```
cgattcagcccgtttttattttaaaaagttatggttagatgcatgggaagaaacgcatgaa
cgtcgattggttgatgaaatttatcgtgttccaaccgaagatactgatcaagcaaaagca
aaagtacaacttgcgttttgtattgatgtacgatccgaaccgtttagaagacatttagaa
agtgaagggccttttgaaacaatagggattgctgggttctttggtctgccaattcaaaaa
```

5   `gaagtacttgatgaacaatttgcacatccatctttaccagtgatggtagaacctgcatat`
    `cgtattaaagaatatgctgatcaacatgaaatgaaaatttataatcaacaacaacataca`
    `cttacatctatgtttacaactttaaattaatgaaaaacaacgtgttgccaagtttgctt`
    `ttaccagaattaagtggtcctttttttaagtattgcgactatagctaacacaattttccct`
    `aaaaaagcaaacgtattgttcatcgattctcacaaaatggctacgtaaaccaacaggt`

10 `aaattaactattcagcgtgagcaagatgcgtattcaaaactaccaatcggctttacttta`
    `gaagaacagattcaattttccaaaaaagcattacaattaatggacttaacagatgatttt`
    `gcaccacttattgttctatgcggacacggtagtgaatcacataataatccctatcatgct`
    `tcattagagtgtgggcttgcggcggtgcctcgagtgggttcaatgcgaaattattagca`
    `gtaatgtgtaatcaagaaatgttagacgtggattattgatggaaggaatcgacattcca`

15 `agacatacagtgtttatagctgctgaacatcaaacgtcagttgatgagttagagtatatt`
    `tatgttccacctttaactacagaagctcaaatgcctttgacgaacttaagcatgtgatg`
    `ccaaaagtatgttataaagccaatttagaacgtttggcatcgttgccaaatataaataac`
    `actgaccataatcctaatgctgaagcgcatcgtcacgctagcgattggagtgaagttcgt`
    `ccagaatggggtctagcacgaaatgctgaattcattattgggaaacgtcaaatcacccaa`

20 `aatagtaatctagagggacgggcatttcttcataattatgattggacaaaggatgaagac`
    `ggtgagattttaaatacaattatttctgggccagcactagtagcacaatggattaattta`
    `caatactacgcctcaaccgtggcacctcactattatggaagcggtagtaaaacaacgcaa`
    `actgtaacaagtggtgtaggtgtcatgcaaggaaatgctagtgatttaatgtatggctta`
    `ccatggcagtcagtaatgatgaatgacaaagaggcgtatcacgcacctattaggctttta`

25 `attgttattcaagcgccagatgcatatattcaacgtttgttaaaacatcataatcacttt`
    `agacaaaaggttgatcatcaatggataagacttgccagtattgatgaaaataatagttgg`
    `aaagactggtag`

Sequence 2010

30 `VKLPYGVQQDAHEVEDALEFINRVITPLSPISTFAARNPWEGLEDASFDQVARWLKSVRD`
    `VDIYFNASTIHRAISNKEIDLKVFEERLDENRAHYNNRSLSDSDINTYIQRAKNLLTIEE`
    `GYFNTKDNEKLEKWVQTNFKDYKKKEDVIAQSASVFTKEGTRLIDILNAHMIKWSKLYVD`
    `DFQSSWTMPKREKGFYHAWQRLVKHDPLFTKKQRLTLAHLPNQATEAIEYAFQELGVKEE`
    `HRQSYIESHLLSLPGWAGIMYHRSQTQSNDAYLLTDYVAIRLSIEMVLLNDHHTTLLKKS`

35 `IYLQKKLEQIRYLLFNIQMNVEQWLNLSSKKQQAYIELGTRFSPFYFKKLWLDAWEETHE`
    `RRLVDEIYRVPTEDTDQAKAKVQLAFCIDVRSEPFRRHLESEGPFETIGIAGFFGLPIQK`
    `EVLDEQFAHPSLPVMVEPAYRIKEYADQHEMKIYNQQQHTLTSMFYNFKLMKNNVLPSLL`
    `LPELSGPFLSIATIANTIFPKKAKRIVHRFSQKWLRKPTGKLTIQREQDAYSKLPIGFTL`
    `EEQIQFSKKALQLMDLTDDFAPLIVLCGHGSESHNNPYHASLECGACGGASSGFNAKLLA`

40 `VMCNQENVRRGLLMEGIDIPRHTVFIAAEHQTSVDELEYIYVPPLTTEAQNAFDELKHVM`
    `PKVCYKANLERLASLPNINNTDHNPNAEAHRHASDWSEVRPEWGLARNAEFIIGKRQITQ`
    `NSNLEGRAFLHNYDWTKDEDGEILNTIISGPALVAQWINLQYYASTVAPHYYGSGSKTTQ`
    `TVTSGVGVMQGNASDLMYGLPWQSVMMNDKEAYHAPIRLLIVIQAPDAYIQRLLKHHNHF`
    `RQKVDHQWIRLASIDENNSWKDW*`

45

Sequence 2011
Contig_0693_pos_971_603,
putative peptide of unknown function
`atgaaaaagacgaaaggtatttatgaatctgaaattagtaaagccattacacaatgggag`

50 `aaagattttttaggaagaggttctttgtctgtaaaaacagatattcttcgtgatatggtt`
    `attgttagtttacaaggtattttgacacctgctgaatatcgcgtgtgtaaaacgaatgaa`
    `gggctttttaaatattaaacgtacacgttctgaacttgtagagtctggtgaagaggacttg`
    `agtcgcattattaaagatttgactggacttaatgtgaaaagttttcatagtgatttaagt`
    `actattaccggtgaacgcgtaatgattttttaagttggaagatcgttttgataaagcatta`

55 `catgagtaa`

Sequence 2012
`MKKTKGIYESEISKAITQWEKDFLGRGSLSVKTDILRDMVIVSLQGILTPAEYRVCKTNE`
`GLLNIKRTRSELVESGEEDLSRIIKDLTGLNVKSFHSDLSTITGERVMIFKLEDRFDKAL`

501

HE*

Sequence 2013
Contig_0694_pos_1943_2533,
putative peptide of unknown function
atgatacaacgtaaaggtgaattaatactttcatggattggcaatggattacatctgttg
tatgtatttttaataggcatattttttcataatgacacaaacaagtgattttaaaaatggg
atgattcaaggatttatagaagaaaatccgggggaatatgatttagcataccaaacttat
aacttgatgttaggtttaggtgttgtcctcattattatacttctcattttattaattgta
tcaatagttgccgctattttaattggcaaaaatgccaaagtatcaggaatattacttgtt
attacaggaatcattggtctcttttttaagtttattgctggtgccttatggttaatagca
ggtatcatgcttttagtacgtaaaccacaaacacaaatgaccaaatcaattctcaatat
agtaatgacatacattcacatgttgtgccagaagaaaagaaacgtgaacaacaacaatat
aatatgaatgaaccacatattggtcaaacatcaacatctcatcatgatcacgcattgaat
gatcaaaataaacgagaaaaccataatcatgataatcaaccatacaaatag

Sequence 2014
MIQRKGELILSWIGNGLHLLYVFLIGIFFIMTQTSDFKNGMIQGFIEENPGEYDLAYQTY
NLMLGLGVVLIIILLILLIVSIVAAILIGKNAKVSGILLVITGIIGLFLSFIAGALWLIA
GIMLLVRKPQTQNDQINSQYSNDIHSHVVPEEKKREQQQYNMNEPHIGQTSTSHHDHALN
DQNKRENHNHDNQPYK*

Sequence 2015
Contig_0694_pos_4776_6005,
putative peptide of unknown function
atgcaaacagtcggaattataccttcgccaggtatagcacatcaacatgcaaaaaaaata
attccaaatgttaaacagttattgtcaaagcgtactaaacatagtcaatggaatttcgac
atcaaagtcgatctcatgataggatctgcagaggatgtacatgaaagtgttgaaaaagca
gcacaaattaaagaggaacatcagtgggattacgttgtttgtctgacagatttgcctagt
atttcagataataaagtggttgtcagcgactttaatagtgacaaacatgttgcaatgcta
tcattaccgtcactaggttttattgatttgaagcgcaagctagttaaaacgatgacttca
ttgattgaacaattatattataatcaaccgaaagacaaaaatgcgccacatcctttgta
cgcgtgaaggctgtagaacctgacgaagacgccacatcaaaacaacgatatattaatatt
ttatttatcataagttggattcagttaattggtggactgacacgagcaaatcagccttgg
aaaaacatctttaattttaagaaaatcatttcagttgcctttgcaacaggaacttatgtc
tcaatattttcaatgccatgggaattaagcgtgatttattcaccgcttcgacttatcata
ttgatggtgattgctatacttgggatggctggatggctattctatgcgcatcaattgatt
gaaaagaaaactgctaaatctcagcgtgtatatcgatatatttataattcaaccacactt
gttacactaagtttgattacactcataaattatgtcattttatatttattgttaatcatc
agtattacactctttgtccctgtggaattatttaatagttggacgagtgcccaatcacaa
tttacgttctcaaattatatgagattgatttggtttgtatcatcattaggactttttagct
ggag ꞏꞏ ꞏꞏꞏ gggatcaactgttgaaaatgaagagaaaatacgtcgtattacttattcꞏtat
agacaatatcatcgttataaagaagctggcaagaacaaaaagaacaagaaacttctcgtg
atgtatcacaacaaatgtcgaacaacaaacttcaagtaaagatgaaaataatgaacaat
atgaaggtaaaaaacaaggacatagagaggaggatgacgcatgacaaatcaaaaaactgt
gggtctagtcgtcgctccaggtgttactga

Sequence 2016
MQTVGIIPSPGIAHQHAKKIIPNVKQLLSKRTKHSQWNFDIKVDLMIGSAEDVHESVEKA
AQIKEEHQWDYVVCLTDLPSISDNKVVVSDFNSDKHVAMLSLPSLGFIDLKRKLVKTMTS
LIEQLYYNQPKDKNAPHPFVRVKAVEPDEDATSKQRYINILFIISWIQLIGGLTRANQPW
KNIFNFKKIISVAFATGTYVSIFSMPWELSVIYSPLRLIILMVIAILGMAGWLFYAHQLI
EKKTAKSQRVYRYIYNSTTLVTLSLITLINYVILYLLLIISITLFVPVELFNSWTSAQSQ
FTFSNYMRLIWFVSSLGLLAGAMGSTVENEEKIRRITYSYRQYHRYKEAGKNKKNKKLLV
MYHNKMSNNKLQVKMKIMNNMKVKNKDIERRMTHDKSKNCGSSRRSRCY*

Sequence 2017
Contig_0696_pos_5640_5990,
is similar to (with p-value 1.0e-27)

>gp:gp|AB015981|AB015981_2 Staphylococcus aureus genes for O
rfA, MnhA, MnhB, MnhC, MnhD, MnhE, MnhF and MnhG, complete c
ds. NID: g4001723.
atgccagtgataagcaaactgaataataatccgagaccatccacatataaattaaaattc
attccaatatgtggcatccaatttaagttttttcattacgttattacctgacatcgtcatt
gaaatcaatgatatgaaatatataaataagacaataggaacaggtaatacaaaccatcct
aaatgtatacgtttaacgaaccggtataggaatggaataatgagtgcaaatattaacggt
aatagtaccgcaatatgtaacaaactcactatgttttcctcctttaaaatttatttacga
cacatattatacatcattgcgcttcttctgaaaaaccatgtaactgcttga

Sequence 2018
MPVISKLNNNPRPSTYKLKFIPICGIQFKFFITLLPDIVIEINDMKYINKTIGTGNTNHP
KCIRLTNRYRNGIMSANINGNSTAICNKLTMFSSFKIYLRHILYIIALLLKNHVTA*

Sequence 2019
Contig_0696_pos_5913_3508,
is similar to (with p-value 0.0e+00)
>gp:gp|AB015981|AB015981_2 Staphylococcus aureus genes for O
rfA, MnhA, MnhB, MnhC, MnhD, MnhE, MnhF and MnhG, complete c
ds. NID: g4001723.
gtgagtttgttacatattgcggtactattaccgttaatatttgcactcattattccattc
ctataccggttcgttaaacgtatacatttaggatggtttgtattacctgttcctattgtc
ttatttatatatttcatatcattgatttcaatgacgatgtcaggtaataacgtaatgaaa
aacttaaattggatgccacatattggaatgaatttttaatttatatgtggatggtctcgga
ttattattcagtttgcttatcactggcataggaagtctagtggttttatattctatcgga
tatttaagcaaatcagaacaactcgggaattttttactgttatttattactatttatgggt
gcaatgttaggagttgtactctctgataattttattattttatatttattttgggagtta
acttctttctctagcttttattaatctctttctggagagagaaaaaagcttcaatttat
ggtgctcaaaaatctttaatcattaccgtattaggtggcttgagcatgctaggtggtatc
attttactttccctagctactgatactttagtattcaggctatgatttcaaaagcaagt
gacattcaaaatagtcctttcttatcttagtaatgatacttttatgattggtgcatttt
acaaatctgcacaagtgcctttttatatttggttaccagatgctatggaggcgcctacg
ccagtgagtgcataccttcattctgcaacgatggtaaaagcaggactatatctaatcgca
agaatcacacctattttgcaatatccgaaggttgggtatggacaattacacttgttggt
ttaatcaccctattttgggcatcactcaatgcaacaaaacaacatgacttaaaaggtatt
ttagctttctcaactgtgtctcaactagggatgattatgtctatgcttggtattggtgct
gtaagttatcattatcaaggcgctaatagtcaactttatgttgctggatttgttgctgcc
atatttcatttaattaatcatgccacgtttaaaggtgcactatttatgattacaggtggt
attgatcattcaactggtacacgtgatgttaaaaagttgggcggtttacttacaatcatg
cctatctcattcacgcttacagttattacaacattaagtatggctggtgtgccgccttt
aacggcttttatcaaaagagaaattcttagagtcaatgattaatgttacacatttaaat
ttaatgagtttaaatactttaggtattctttttaccaatcattgccattattggtagtatt
ttcacatttgtatattcaattaaatttatattgcatatattctttggttcttataaacct
gaagctctgccaaaacaagcgcatgaatcttcaattttaatgcttatttcacctatcatt
ttaacatcactagttatagtattcggtttattcccaagtatattaacgcaatctattata
gagccggcatctgtagcagttagtcaaacatcaaatataactgctgagttccatttattc
catggtataactccagcattcctatcaacaataggtatttatattattggtattttatta
ttaatttcatttagttattgggttcgtttattacaagcacatccatatcagttaacgttg
aatcattggtatgacacttcaggccaacgtattccaggatattccgaaaatataacaaat
agttatgttacaggttttttctagaaataaatttggtgattatcttaggtattctcattgct
ttaacttttgttacagtcatcagtgtacccttcagtattgactttaaaaacgtgagtcat
ttacgcgtatttgaaggtgcaacagtattgttttttactgattgcttcaactttcattata
tttgcaaaatcacgtttgtttagcattatcatgttaagcgctgtgggttacgctattca
gtattatttattttctttaaagcgcccgacttagcattaacacaatttgtagtggaatct
atttctacagcattattcttactatgctttttatcacctacctaatttaaatcgctacaat
gaaaaaccaacctttaaactgacaaatgctgtgatttcaattggagtgggattatcagtg
attattttaggattaattggctatggtaatagacactttgactctattactaaattctat
caagaacatgtttttgatttagcacatggtaaaaaatatggtaaatgtcatcctcgtagat
ttccgtggtatggatactttattcgagtcatctgtactaggtattgcaggtttaggcgta

tatacaatgattaaattacgattaaaacagaaaaatcaatcaagtgaggtgaatgaccat
gaatag

Sequence 2020
VSLLHIAVLLPLIFALIIPFLYRFVKRIHLGWFVLPVPIVLFIYFISLISMTMSGNNVMK
NLNWMPHIGMNFNLYVDGLGLLFSLLITGIGSLVVLYSIGYLSKSEQLGNFYCYLLLFMG
AMLGVVLSDNFIILYLFWELTSFSSFLLISFWREKKASIYGAQKSLIITVLGGLSMLGGI
ILLSLATDTFSIQAMISKASDIQNSPFFILVMILFMIGAFTKSAQVPFYIWLPDAMEAPT
PVSAYLHSATMVKAGLYLIARITPIFAISEGWVWTITLVGLITLFWASLNATKQHDLKGI
LAFSTVSQLGMIMSMLGIGAVSYHYQGANSQLYVAGFVAAIFHLINHATFKGALFMITGG
IDHSTGTRDVKKLGGLLTIMPISFTLTVITTLSMAGVPPFNGFLSKEKFLESMINVTHLN
LMSLNTLGILLPIIAIIGSIFTFVYSIKFILHIFFGSYKPEALPKQAHESSILMLISPII
LTSLVIVFGLFPSILTQSIIEPASVAVSQTSNITAEFHLFHGITPAFLSTIGIYIIGILL
LISFSYWVRLLQAHPYQLTLNHWYDTSGQRIPGYSENITNSYVTGFSRNNLVIILGILIA
LTFVTVISVPFSIDFKNVSHLRVFEGATVLFLLIASTFIIFAKSRLFSIIMLSAVGYAIS
VLFIFFKAPDLALTQFVVESISTALFLLCFYHLPNLNRYNEKPTFKLTNAVISIGVGLSV
IILGLIGYGNRHFDSITKFYQEHVFDLAHGKNMVNVILVDFRGMDTLFESSVLGIAGLGV
YTMIKLRLKQKNQSSEVNDHE*

Sequence 2021
Contig_0696_pos_3473_3087,
is similar to (with p-value 4.0e-57)
>gp:gp|AB015981|AB015981_3 Staphylococcus aureus genes for O
rfA, MnhA, MnhB, MnhC, MnhD, MnhE, MnhF and MnhG, complete c
ds. NID: g4001723.
gtgattattttcttcatggtcattgtgtttggcttttcattattttttagctggacattat
acacctggtggcggtttcgttggtggtttgcttttcgctagtgcattattagttattaca
attgcatatgatgtaaaaacgatgcgaaagattttttcctttagattttaaaatcttaatt
ggtattggtttattattttgcgtgggtacaccattaacaagttggttcatgtctaaaaac
ttttttacacatgtcacttttgacatccctttgcctttacttgaacctatgcacatgacg
acagcgatgttttttgatttcggtgtttttatgtgcagttgtaggaactattatgactata
attatttcgattggagagaacgaatag

Sequence 2022
VIIFFMVIVFGFSLFLAGHYTPGGGFVGGLLFASALLVITIAYDVKTMRKIFPLDFKILI
GIGLLFCVGTPLTSWFMSKNFFTHVTFDIPLPLLEPMHMTTAMFFDFGVLCAVVGTIMTI
IISIGENE*

Sequence 2023
Contig_0696_pos_3075_2740,
is similar to (with p-value 2.0e-37)
>gp:gp|AB015981|AB015981_4 Staphylococcus aureus genes for O
rfA, MnhA, MnhB, MnhC, MnhD, MnhE, MnhF and MnhG, complete c
ds. NID: g4001723.
atgatctttgttagtggtattctcacatctataagtgtctatctcgttttgtctaaaagt
ttgatacgtatcattatggggactacactactaactcatgctgcaaatttattttttaatt
actatgggaggtttaaagcacggaactgttccaatatttgaaaaaggaacatcaagctat
gttgaccctatcccccaagcattgattttaacagctatcgttatcgcctttgctacaaca
gctttcttttttagttcttgcatttagaacatataaagaactaggcactgataacgttgag
ctaatgaaaggagcgccagaagatgatagagagtaa

Sequence 2024
MIFVSGILTSISVYLVLSKSLIRIIMGTTLLTHAANLFLITMGGLKHGTVPIFEKGTSSY
VDPIPQALILTAIVIAFATTAFFLVLAFRTYKELGTDNVELMKGAPEDDRE*

Sequence 2025
Contig_0696_pos_2408_1257,
is similar to (with p-value 0.0e+00)
>gp:gp|AB015981|AB015981_5 Staphylococcus aureus genes for O

504

rfA, MnhA, MnhB, MnhC, MnhD, MnhE, MnhF and MnhG, complete c
ds. NID: g4001723.
atgtttatgttgattggtattattggctcatttacaacaggagatattttcaacttgttt
gtgttctttgaagtctttttaatgtcttcatattgtttactcgttattggtactactaaa
atacaattacaagaaacaattaagtatattttagtcaatgttgtttcatcgtctttctttt
gtcatgggtgttgcagtttttatattcagttgtaggaactttaaatctcgctcatattagt
gaaagattgtcacaactttctgtacatgacagtggcttagtcaatattgttttttattttta
tttatctttgtctttgccactaaagcaggcgttttcctatgtacgtatggctacctggt
gcttattatgcccctccagtagcgatcatcacgttctttggtgcactattgactaaagtg
ggtgtatacgcaattgcgagaactctaagtttattctttaataatacagtaagcttttct
cattatgtcatcctttcttagcattacttacaattattttttggatgtataggtgcgata
gcttactatgatacgaagaaaatcatcctttacaatattatgattgcagtaggtgtcata
ttagttggtattgctatgatgaacgaatcaggcatgactggtgcaatatattacacacta
catgatatgttagttaaagcttcattgttcttactcattggcgtcatgtacaaaatcact
aaaacgactgacttacgtcattttggtggcttgataaaagggtatcctattctaggttgg
ac?ttctttattgcagcgctaagcttagcgggtataccacctttt agtggtttcta_ggt
aaat?ct?atattgttcgagcgacctttgaaaaaggatttttatctaagtggtatca tgta
cttttatcaagtttaatcgtgttatattcagtcatacgtatttttcttaaaaggattttt c
ggtgaagttgaaggatatactttatctaaaaaggtaaatgttaaatatctaacaactatc
gctgttgcatctacagttattactgtaatctttggattatctgcagacacgttattccca
atcatcaaagatggcgctgaaacgtttgtcgatccaagtcaatatattcatagtgtgtta
ggaggtaaatag

Sequence 2026
MFMLIGIIGSFTTGDIFNLFVFFEVFLMSSYCLLVIGTTKIQLQETIKYILVNVVSSSFF
VMGVAVLYSVVGTLNLAHISERLSQLSVHDSGLVNIVFILFIFVFATKAGVFPMYVWLPG
AYYAPPVAIITFFGALLTKVGVYAIARTLSLFFNNTVSFSHYVILFLALLTIIFGCIGAI
AYYDTKKIILYNIMIAVGVILVGIAMMNESGMTGAIYYTLHDMLVKASLFLLIGVMYKIT
KTTDLRHFGGLIKGYPILGWTFFIAALSLAGIPPFSGFYGKFYIVRATFEKGFYLSGIIV
LLSSLIVLYSVIRIFLKGFFGEVEGYTLSKKVNVKYLTTIAVASTVITVIFGLSADTLFP
IIKDGAETFVDPSQYIHSVLGGK*

Sequence 2027
Contig_0696_pos_1081_776,
is similar to (with p-value 3.0e-40)
>gp:gp|AB015981|AB015981_6 Staphylococcus aureus genes for O
rfA, MnhA, MnhB, MnhC, MnhD, MnhE, MnhF and MnhG, complete c
ds. NID: g4001723.
atgcttatcattacattttttaactgagttaataaaagcaaactttggtgtactaaaaatt
attctcaaaccacgaattgagaataaacccggattctttgtgtacgagacggaattagaa
cgtgactggcaacttgttttactttccaacttgattacgttaacacctggcacagtcgtt
ttaggtattagtgatgaccgtaaaaagatttatatccactcaattgatttcagtacaaag
gaagaagagattcaaaatatcaaatcttcattagagaaggtcgttagaaaggtaggcgag
aaataa

Sequence 2028
MLIITFLTELIKANFGVLKIILKPRIENKPGFFVYETELERDWQLVLLSNLITLTPGTVV
LGISDDRKKIYIHSIDFSTKEEEIQNIKSSLEKVVRKVGEK*

Sequence 2029
Contig_0697_pos_2646_3434,
is similar to (with p-value 3.0e-34)
>sp:sp|P06696|TNPA_STAAU TRANSPOSASE A (TRANSPOSON TN554). >
pir:pir|A24584|A24584 transposition regulatory protein tnpA
- Staphylococcus aureus transposon Tn554 >gp:gp|X03216|ISTN5
54_i Staphylococcus aureus transposon Tn554. NID: g43726. >g
p:gp|K02987|TRN554_1 Transposon Tn554 (from S.aureus), compl
ete, containing transposition genes tnpA, tnpB and tnpC, and
antibiotic resistance genes ermA and spc. NID: g154920.

```
gtgtatacttatactattaaaattagaaatgagattattatgaaaatagtagaagtaaaa
tctaagaatggtaccaattttatgattttagatggtaataatgaacctatagtagatgca
gtaagatatttgaagtatctggatagtgttaagaaaagtttaaataccaagaaaacctat
gcctatgcactaaaaaatttttttgtttacttagaaagtaaaaagatatgctataaagaa
gttagttttgataactttgttgattttataagatggatgaaaacacctttgaatatgag
aatgtcctctcttatcaccgaaaagaaaaaagcattagtcctaagacaattaatctgact
atgactgtagtatctaattttttatgattatctctataggagtaaaaaattagatgttaat
ttctatgattttatgcatatggaaagtaaatactctaaaaaatataaaagtttcatgcat
cacataaataaggactatagaacgttgaaaaatattttgaaagttaaagaaccaaagaaa
aaaatagaagtgttaactaatgcggaggttaagaaattattagaggaagctaataatatt
agagataaattcttaatacaattactatatgaaaccggattacgtataggtgaggtatta
tcattacgtattgatgatattaaatttgactttagaaacccatcgattaaaacgatacta
ggtggtacaggaatatcaacctgttatccatcgcctacgcctgtcggcctcagcttagga
cccgactaa
```

Sequence 2030
VYTYTIKIRNEIIMKIVEVKSKNGTNFMILDGNNEPIVDAVRYLKYLDSVKKSLNTKKTY
AYALKNFFVYLESKKICYKEVSFDNFVDFIRWMKTPFEYENVLSYHRKEKSISPKTINLT
MTVVSNFYDYLYRSKKLDVNFYDFMHMESKYSKKYKSFMHHINKDYRTLKNILKVKEPKK
KIEVLTNAEVKKLLEEANNIRDKFLIQLLYETGLRIGEVLSLRIDDIKFDFRNPSIKTIL
GGTGISTCYPSPTPVGLSLGPD*

Sequence 2031
Contig_0697_pos_5460_5116,
putative peptide of unknown function
```
gtgacaaaccggaggaaggtggggatgacgtcaaatcatcatgccccttatgatttgggc
tacacacgtgctacaatggacaatacaaagggcagcgaaactgcgaggtcaagcaaatcc
cataaagttgttctcagttcggattgtagtctgcaactcgactatatgaagctggaatcg
ctagtaatcgtagatcagcatgctacggtgaatacgttcccgggtcttgtacacaccgcc
cgtcacaccacgagagtttgtaacacccgaagccggtggagtaaccatttggagctagcc
gtcgaaggtgggacaaatgattggggtgaagtcgtaacaaggtag
```

Sequence 2032
VTNRRKVGMTSNHHAPYDLGYTRATMDNTKGSETARSSKSHKVVLSSDCSLQLDYMKLES
LVIVDQHATVNTFPGLVHTARHTTRVCNTRSRWSNHLELAVEGGTNDWGEVVTR*

Sequence 2033
Contig_0697_pos_1439_399,
putative peptide of unknown function
```
atgcaaaaagtacgttctgatataatgactcatcgtggttctcactatgatttaggagta
aagactgctctttggttacaaaccactcctttattaaaaaatcgaaataaagaatggcga
aagagaattccacgttttgatattgatgtcaaggaaacctatgatatatttcaaatctat
tcgccacagatttgggaagaaataattggtatgcaagatgtattgaatctacctacaaaa
caaatgatttttaaattttggccattatcgatttactgatttaaaggacagtggttgcaca
gtatataaaggtcgtgatttttttagtccgaaattatgattatcatcctgcaacatatgat
ggtagatacttattattttcaacctaatgacgggggattatctcaaataggaccgacttca
agagtgactggtagaatggatggtatgaacgagtatggtttagttatggcatataatttt
atgcatcgtaaaaagcctgcaaatggatttgtatgttacatggttggtcggctaaaactt
gaaaattgcaaaaatgtaactgaagcaatcaaattttttaaaggaagtaccgcatcgtagt
tcattcagttatatactaatggatagacattcgaattatgccattgtcgaagttacacct
cgatcaatagatgtaaggtatgaacatatatgcacaaatcattttgaattgcttacccat
gaaaatagaaactatacaagagaatctaaagaacgcttaaatcgtgtaataaataaaaca
actccttctacaaacaaagatatcgcattcaaattatttaacgacccgcaatacgaaatc
tatagcaacctatttaaaagttggtctggtacaattcatacttcactatatgaacctaat
tcattaatatcatggatggcattaggtcaaaacagtcatccaacctcaatcaatttttct
aattggttaaaaggaaagaaattgaatataaattactttgaaggcgaaatagatacacca
ttaacttttgccacatactaa
```

Sequence 2034

MQKVRSDIMTHRGSHYDLGVKTALWLQTTPLLKNRNKEWRKRIPRFDIDVKETYDIFQIY
SPQIWEEIIGMQDVLNLPTKQMILNFGHYRFTDLKDSGCTVYKGRDFLVRNYDYHPATYD
GRYLLFQPNDGGLSQIGPTSRVTGRMDGMNEYGLVMAYNFMHRKKPANGFVCYMVGRLIL
ENCKNVTEAIKFLKEVPHRSSFSYILMDRHSNYAIVEVTPRSIDVRYEHICTNHFELLTH
5   ENRNYTRESKERLNRVINKTTPSTNKDIAFKLFNDPQYEIYSNLFKSWSGTIHTSLYEPN
SLISWMALGQNSHPTSINFSNWLKGKKLNINYFEGEIDTPLTFATY*

Sequence 2035
Contig_0699_pos_1065_1427,
10   is similar to (with p-value 1.0e-26)
>gp:gp|AB003188|AB003188_3 Micrococcus luteus hexs-a, menG,
hexs-b gene, complete cds. NID: g2982678.
gtggcaaagttaaacattaacaacgaaataaagaaagtagaaaagcgacttgaagaagca
attataagttctgatcaaacattacaagaagcctcattccatttactatcttcaggggga
15   aaaagagttagacccgcttttgttattttaagtggtcaatttggctctaacaacaaacct
tcagaagacacgtatcgtgtagcagtagctttagaactaattcacatggctaccttagtc
cacgatgatgtgatagataaaagtgataaacgtagaggccgactcactatttcaaaaaaa
tgggaccaaagtacagctattttaacaggaaatttcttacttgctatggggctcaagcac
tga
20

Sequence 2036
VAKLNINNEIKKVEKRLEEAIISSDQTLQEASFHLLSSGGKRVRPAFVILSGQFGSNNKP
SEDTYRVAVALELIHMATLVHDDVIDKSDKRRGRLTISKKWDQSTAILTGNFLLAMGLKH
*
25

Sequence 2037
Contig_0699_pos_2988_3362,
is similar to (with p-value 1.0e-18)
>sp:sp|P31114|GRC3_BACSU PROBABLE HEPTAPRENYL DIPHOSPHATE SY
30   NTHASE COMPONENT II (EC 2.5.1.30) (HEPPP SYNTHASE) (SPORE GE
RMINATION PROTEIN C3). >gp:gp|M80245|BACVARGNS_5 B.subtilis
dbpA, mtr(A,B), gerC(1-3), ndk, cheR, aro(B,E,F,H), trp(A-F)
, hisH, and tyrA genes, complete cds. NID: g143798. >gp:gp|Z
99115|BSUB0012_214 Bacillus subtilis complete genome (sectio
35   n 12 of 21): from 2195541 to 2409220. NID: g2634478.
atgatcggacattatataggtatgagtttccaaataatagatgatgtgctagattttact
agttctgaaaagaaacttggtaagccggttggtagtgacccttatgaatggtcatattaca
ttacctgtactattagaaatgcgaaaaaataagacttttaaagataaaatttcacaactt
aatcctgacagtcctcaacatgcctttgaaacttgtataacaataattagacagtccgaa
40   agcatagaacaatcaaaacaaataagtgaaaagtatttaaataaagcaatcaatttaatc
gatgaattagaggatggtcctaataaagaactatttagaaagcttattaaaaaaatggga
agtcgaaataagtaa

Sequence 2038
45   MIGHYIGMSFQIIDDVLDFTSSEKKLGKPVGSDLMNGHITLPVLLEMRKNKTFKDKISQL
NPDSPQHAFETCITIIRQSESIEQSKQISEKYLNKAINLIDELEDGPNKELFRKLIKKMG
SRNK*

Sequence 2039
50   Contig_0699_pos_4537_5598,
is similar to (with p-value 0.0e+00)
>sp:sp|Q59803|AROC_STAAU CHORISMATE SYNTHASE (EC 4.6.1.4) (5
-ENOLPYRUVYLSHIKIMATE-3-PHOSPHATE PHOSPHOLYASE). >gp:gp|U319
79|SAU31979_3 Staphylococcus aureus chorismate synthase (aro
55   C) and nucleoside diphosphate kinase (ndk) genes, complete c
ds, dehydroauinate synthase (aroB) and geranylgeranyl pyroph
osphate synthetase homolog (gerCC) genes, partial cds. NID:
g987495.
atgtttaaacgtcaaggggggatatggtcggggggcgtcgtatgaaaattgaaaaagaact

```
atagaaattgtgtcaggcgtcagaaatggctttacttttaggaagtccaataactttagta
gttactaacgacgatttcactcattggagaaaaataatgggcgtcgcaccaattagtgat
gaagaaagagaaaatatgaaacgtaccattacaaaacctagaccgggccatgctgatctt
ataggtggcatgaaatataatcatcgtgatttaagaaatgtgcttgagcgttcatctgct
```
5
```
agagaaacagcagcaagagttgctgtgggtgctgtttcaaaaattcttttagaacaatta
gatatccacttatatagccgtgtagtcgaaattggtggtattaaagacaaaggtttatat
gatgtagatatgttcaaaaataatgtagataaaaatgatgtacgtgtaattgacgaaaat
attgcgcaacaaatgagagataaaatagatgaagcgaaaaaagacggagattcaatcggg
ggcgtagttcaagtaatggctgaaaacatgcctattggagtgggaagttatgtacactat
```
10
```
gaccgtaaattagatggacgcattgcacagggtgttgtgagtatcaacgccttcaaaggt
gtaagttttggtgagggatttaaagcagctgaaaaacctggtagcgaaattcaagatgaa
attcattataatcaagattcaggctatttttagagctacaaatcacttaggtggatttgaa
ggaggcatgagtaatgggatgcctataattgttaatggtgtcatgaagcctattcctact
ttatataaaccactaaactcagttgatattaatactaaagaagacttcaaagctactata
```
15
```
gaacgctcagatagttgtgcagtgcccgcagctagcgtagtatgtgaacacgttgtcgct
tttgagttagcaaaagcagtactcgaagagtttcaatctaaccacatggaccaactcgta
gcacaaattaaagagcgtcgacaactcaacatagaatttttaa
```

Sequence 2040

20
```
MFKRQGGYGRGRRMKIEKDTIEIVSGVRNGFTLGSPITLVVTNDDFTHWRKIMGVAPISD
EERENMKRTITKPRPGHADLIGGMKYNHRDLRNVLERSSARETAARVAVGAVSKILLEQL
DIHLYSRVVEIGGIKDKGLYDVDMFKNNVDKNDVRVIDENIAQQMRDKIDEAKKDGDSIG
GVVQVMAENMPIGVGSYVHYDRKLDGRIAQGVVSINAFKGVSFGEGFKAAEKPGSEIQDE
IHYNQDSGYFRATNHLGGFEGGMSNGMPIIVNGVMKPIPTLYKPLNSVDINTKEDFKATI
```
25
```
ERSDSCAVPAASVVCEHVVAFELAKAVLEEFQSNHMDQLVAQIKERRQLNIEF*
```

Sequence 2041
Contig_0701_pos_1174_2742,
is similar to (with p-value 0.0e+00)

30
```
>gp:gp|Z99111|BSUB0008_149 Bacillus subtilis complete genome
(section 8 of 21): from 1394791 to 1603020. NID: g2633699.
>gp:gp|Z97025|BSZ97025_8 Bacillus subtilis nprE, yla[A,B,C,D
,E,F,G,H,I,J,K,L,M,N,O] and pycA genes. NID: g2224758.
atggttgacggtgtcgtactagtggttgacgcatatgaaggtacaatgcctcaaactcgt
```
35
```
tttgttcttaaaaaagctttagaacaaaacttaaaaccggttgtagttgtgaataaaatt
gataaaccagctgctagacctgagggagttgtagatgaagtattagacttattcattgaa
ttggaagcgaatgatgagcaattagacttcccagttgtttatgcttcagctgtgaatgga
acagcaagtttagactctgaaaagcaagacgaaaatatgcaatccctatacgagacgatt
attgactatgtaccggcaccagtagataattcagatgaaccattacaattccaaattgct
```
40
```
ttactagattataatgattatgtaggtcgtataggcgttggacgtgtgttcagaggtaaa
atgcgtgtaggtgataatgtatcactaattaaattagatggtacagttaagaactttcgt
gtgacgaaaatatttggttactttggtcttaaacgtgaagaaattgaagaagcacaagca
ggagacttaatagctgtttcaggtatggaagatattaacgttggtgaaacagttacacca
catgatcatcgtgacccattaccggtgttacgtattgatgaaccaaccctagaaatgact
```
45
```
tttaaagtaaataactctccgtttgctggacgtgaaggtgattatgtaacagctcgacaa
attcaagaaagattagatcaacaacttgaaacagatgtttctttaaaagttacacctact
gatcaaccagattcatgggttgttgctggtcgtggtgaactacacttgtctattcttatt
gaaaacatgagacgtgaaggctttgaattacaggtttctaaacctcaagttattttaaga
gaaatcgatggtgtgttaagtgaaccatttgagcgtgtacaatgtgaagtgccttctgaa
```
50
```
aatgccgggcagtgattgagtcattaggtgcacgaaaaggtgaaatgttagatatgatg
acgaccgacaatggtttgacgcgtttaatctttatggtacctgcacgcggtatgattggt
tatactactgaatttatgtctatgacacgaggttatggaattattaaccatacatttgaa
gaatttagacctcgcgttaaagctcaaatcggtggtagacgtaacggtgcattgatttct
atggaccaaggtcaagcaacatcttatgcgattattaacttagaagatcgtggtgttaac
```
55
```
tttatggaaccaggtactgaagtatatgaaggtatgattgttggtgaacataaccgtgag
aacgatttaacagtaaatattactaaagcaaagcatcaaacaaacgtacgttcagctact
aaagatcaaacacaaacgatgaatcgtcctagaattttaacattagaagaagcgttacaa
tttatcaatgatgatgaattggtggaagtaactcctgaaagtattcgcttaagaaagaaa
atacttaataaatctgcccgtgaaaaagaagcaaaaagagttaaacaattaatgcaagac
```
```

gaacaataa

Sequence 2042
MVDGVVLVVDAYEGTMPQTRFVLKKALEQNLKPVVVVNKIDKPAARPEGVVDEVLDLFIE
5   LEANDEQLDFPVVYASAVNGTASLDSEKQDENMQSLYETIIDYVPAPVDNSDEPLQFQIA
LLDYNDYVGRIGVGRVFRGKMRVGDNVSLIKLDGTVKNFRVTKIFGYFGLKREEIEEAQA
GDLIAVSGMEDINVGETVTPHDHRDPLPVLRIDEPTLEMTFKVNNSPFAGREGDYVTARQ
IQERLDQQLETDVSLKVTPTDQPDSWVVAGRGELHLSILIENMRREGFELQVSKPQVILR
EIDGVLSEPFERVQCEVPSENAGAVIESLGARKGEMLDMMTTDNGLTRLIFMVPARGMIG
10  YTTEFMSMTRGYGIINHTFEEFRPRVKAQIGGRRNGALISMDQGQATSYAIINLEDRGVN
FMEPGTEVYEGMIVGEHNRENDLTVNITKAKHQTNVRSATKDQTQTMNRPRILTLEEALQ
FINDDELVEVTPESIRLRKKILNKSAREKEAKRVKQLMQDEQ*

Sequence 2043
15  Contig_0701_pos_3196_4209,
putative peptide of unknown function
atggaacgattttgttgtgtaaatcaaattaactatattcaaatgaatccgttagaagcc
aaatttaaaacgagcgctctaagatcatggaaaactgatcaggcagatgctcataagctt
gcttgtttaggaccgacgcttaaacaaacagacagcttacctatacatgagttaatattc
20  tttgaattaagagaacgcgtccgttttcatctagaaatcgagaatgaacaaaatcgactt
aaatttcagatccttgaattactccatcaaacattccctggtttagaaagattgtttagt
agtcgatattcaatcattgcactcaacatcgcagaaatctttactcattcagacatggtt
cttgatatcgacaaggaggtactgattacacatatattcaattctacagataagggaatg
tcaatggataaagctacaaaatatgcacttcaattaagggtgattgctcaagaaagctat
25  cctaatgtcgatagacattcctttctagtcgaaaaattacgcttacttattcaacaatta
aaacaatctattcatcatctcaaacaattagatgatgccatgattcaattagcacaacaa
ctcgattattttgaaaatattcattcgatacctggtattggtaagctaagcacagctatg
attattggggagattggtgatattaagcgatttaaatcaaataaacaactcaatgctttt
gttggcattgatatcaaacgatatcaatcaggtcatacacactgtagagataccatcaac
30  aagcgtggtaataaaaaagcgagaaaaacttttattttgggtgattatgaatataataaga
gggcagcatcattatgacaatcatgtcgtcgattattactacaaactaagaaagcagcct
aatgagaaacctcataagactgccatcattgcttgtataaatcgattattaaaaacaatt
cattatcttgtaatgaatcataaattgtacgattatcaaatgtcaccacattag

35  Sequence 2044
MERFCCVNQINYIQMNPLEAKFKTSALRSWKTDQADAHKLACLGPTLKQTDSLPIHELIF
FELRERVRFHLEIENEQNRLKFQILELLHQTFPGLERLFSSRYSIIALNIAEIFTHSDMV
LDIDKEVLITHIFNSTDKGMSMDKATKYALQLRVIAQESYPNVDRHSFLVEKLRLLIQQL
KQSIHHLKQLDDAMIQLAQQLDYFENIHSIPGIGKLSTAMIIGEIGDIKRFKSNKQLNAF
40  VGIDIKRYQSGHTHCRDTINKRGNKKARKLLFWVIMNIIRGQHHYDNHVVDYYYKLRKQP
NEKPHKTAIIACINRLLKTIHYLVMNHKLYDYQMSPH*

Sequence 2045
Contig_0701_pos_6054_6374,
45  putative peptide of unknown function
atgtaccggcgatggtatcttttttcaactacagctacttgcttaccactttgctttaatg
ttaacgcagcatgccaagctgcatgaccacttcctaaaaatactacatcatattgtttca
ttaacactcatcctttcttatttttctatgagatgtttaatgtttgctctagttcttca
aacacatatttcgtttcatcatactcagtcgtatcaaattcttgtggtaagcaacttgaa
50  attgcttcaaaaacagcttcttgttgttgttgcccattgtcagttaacgtaattatcaac
tgtcgtttatcagattgttga

Sequence 2046
MYRRWYLFQLQLLAYHFALMLTQHAKLHDHFLKILHHIVSLTLILSYFSMRCFNVCSSSS
55  NTYFVSSYSVVSNSCGKQLEIASKTASCCCCPLSVNVIINCRLSDC*

Sequence 2047
Contig_0702_pos_13157_13459,
putative peptide of unknown function

```
          gtgttctgtagtaggggcaaaaccaatttgataagttgtttgagttttagttttagttat
          ctttctctctacttttttaggttggagatctacagacttgttttggccatcccttttaat
          ttttacagttgttttcgccgtttttattttggtcaagaacatgtttgatatcatcaaaatt
          tttaattttatgattacctacctgaactattttatcacctttgtgcaatccagcttcatc
 5        agctggagatttcttcacaacttctcctatgacattggttggcgtaccttggtagtatgc
          taa
```

```
          Sequence 2048
          VFCSRGKTNLISCLSFSFSYLSLYFFRLEIYRLVLAIPFNFYSCFRRFILVKNMFDIIKI
10        FNFMITYLNYFITFVQSSFISWRFLHNFSYDIGWRTLVVC*
```

```
          Sequence 2049
          Contig_0702_pos_14035_12749,
          is similar to (with p-value 2.0e-24)
15        >sp:sp|P56136|Y258_HELPY HYPOTHETICAL PROTEIN HP0258. >gp:gp
          |AE000545|HPAE000545_4 Helicobacter pylori section 23 of 134
          of the complete genome. NID: g2313349.
          atgagcattttaattacgattatttcatttatcatcgtatttggtgtactcgtaactgtt
          cacgaatatggacacatgttttttgctaagcgagcaggaattatgtgtcctgaatttgcg
20        attggtatggggtcctaaaattttttagttttcgtaaagatgaaacattatatacaattcgt
          ctattaccggtgggtggttatgtcaggatggctggtgatggtcttgaagaaccaccagtt
          caaccaggtatgaacgtaaaaataaagttaaataaccaagacgaaatcacacatataatt
          ctagatgaccaacataaattccaacaaattgaagccatagaagttaagaaatgtgatttt
          aaagatgacctatatattgaaggtatcacttcttatgatgatgaaaggcatcacttcact
25        atagcgaaaaaggcattttttgtcgaaatggaagccttgttcaaattgctccaagagat
          agacagtttacacataagaaaccattgccaaagtttttaacattatttgcaggtccgtta
          tttaattttattttagctttagttctatttattggattagcatactaccaaggtacgcca
          accaatgtcataggagaagttgtgaagaaatctccagctgatgaagctggattgcacaaa
          ggtgataaaatagttcaggtaggtaatcataaaattaaaaattttgatgatatcaaacat
30        gttcttgaccaaaataaaacggcgaaaacaactgtaaaaattaaaagggatggccaaaac
          aagtctgtagatctccaacctaaaaaagtagagagagaagataactaaaactaaaactcaa
          acaacttatcaaattggttttgcccctactacagaacacagcgtttttaaaccaataagc
          tacgtgtattataacttttttcgataaaggtaagcttattttttacagctgttgttgqtatg
          ttagctagtatatttacaggagaattttcatttgatatgttaaatggccctgttggtatt
35        tatcacagtgttgattctgttgttaaatctggaattattaatttagtaggatacaccgct
          ttattaagtgttaacttaggaataatgaatttgctacctattccagcgcttgatggtggt
          cgcatattatttgtactatatgaggctattttttagaaaaccagtgaataaaaaagcggaa
          acaggaattattgctgtaggcgcacttttttgtggttattattatgattttagtcacttgg
          aatgatatacaacggtatttcttataa
40
```

```
          Sequence 2050
          MSILITIISFIIVFGVLVTVHEYGHMFFAKRAGIMCPEFAIGMGPKIFSFRKDETLYTIR
          LLPVGGYVRMAGDGLEEPPVQPGMNVKIKLNNQDEITHIILDDQHKFQQIEAIEVKKCDF
          KDDLYIEGITSYDDERHHFTIAKKAFFVENGSLVQIAPRDRQFTHKKPLPKFLTLFAGPL
45        FNFILALVLFIGLAYYQGTPTNVIGEVVKKSPADEAGLHKGDKIVQVGNHKIKNFDDIKH
          VLDQNKTAKTTVKIKRDGQNKSVDLQPKKVERKITKTKTQTTYQIGFAPTTEHSVFKPIS
          YGIYNFFDKGKLIFTAVVGMLASIFTGEFSFDMLNGPVGIYHSVDSVVKSGIINLVGYTA
          LLSVNLGIMNLLPIPALDGGRILFVLYEAIFRKPVNKKAETGIIAVGALFVVIIMILVTW
          NDIQRYFL*
50
```

```
          Sequence 2051
          Contig_0702_pos_10795_6467,
          is similar to (with p-value 0.0e+00)
          >gp:gp|D86727|D86727_1 Staphylococcus aureus DNA for DNA pol
55        ymerase III, complete cds. NID: g1483181.
          gtggttattattttggcaatgacaaatcgagaaaagtttaaagtgcttgccgatcaaata
          aaaatatcaaatcaactagaacaagatattcttgaacaaggtgaactcactcgtatagat
          gtttcaaataaaaacagaacatggactttccaaatatcactcccacattttttatctcat
          gaagattatcttctttttacacatgcaattgaagaagaatttaaagaaatagctacagta
```

```
      gcaattgattttttcaattaaagataccaacaatcaagatgagtttgctttaaaatatttc
      ggacattgtattgatcaaacacgattgtcgccaaaagtgaaaggtcaattgaaacaaaaa
      aaactcattatgagtggaaatgttttaaaagtcttagtttcaaatgacattgagagaaat
      cattttgataaggcatgtaatggtagtttggttaaagcatttagacagtgtggctttgaa
  5   attgataaagtcgttttgaaacagattcaacaaatcacgatgatgaccttgcatcgtta
      gaagcacatattcaacaagaagatgaacaaagtgcaagagaagcaactgaaaaattagaa
      aaaatgaaagcagaaaaagcgaaacaacaagataataatgaaagtacagtggaaaaatgt
      cagattggaaaaccaattcagattgaaaatataaaaccaattgaatcaattattgaagaa
      gaattcaaagtagctattgaaggtgttatatttgatattaacctaaaagaacttaaaagt
 10   ggacgtcatatagttgagcttaaagttactgattacacagattcacttgtattaaaaatg
      tttacaagaaaaaataaagatgacttggaccactttaaggcacttagtgttggtaaatgg
      gttagagctcaaggtcgtattgaagaagatacttttgttagggatcttgtcatgatgatg
      tcagatattgaagaaattaaaaagacacctaaacaagataaagcagaagataagcgtgta
      gagtttcatttacatacgtctatgagtcaaatggatggtattcctaatattagtgcatat
 15   gttgaacaagctgctaaatggggggcaccaagctttagcagtaacagatcacaacgtagta
      caagcttttcctgatgcacataatgctgccgaaaaacatggtattaagatgatttatggt
      atggaaggtatgctagtagacgatggtgttcctatagcttataaaccaacagaccgtaat
      ttaaaagatgcaacatatgtggtgtttgacgtagagacaacaggtctttctaatcaatat
      gataaaattattgaattagctgcagtaaaagtgcataacggtgaaattatagataagttt
 20   gaacgttttagtaatccacacgaaagattatctgaaaccattatcaatcttacacatatc
      actgatgatatgttaactgatgctcccgaaattgaagaagtgttaactgaatttaaagag
      tgggttggagatgctatatttgtagctcataatgcttcatttgatatgggatttattgac
      acaggatatgaaaggttaggccttttggaccttctacaaacggtgtaattgatacacttgag
      cttcacgtacaattaataccgaatatgggaaacatggtttgaatttccttgccaaaaaa
 25   tatggtgtcgaattaacgcaacatcatagagcgatttatgatacagaagcaacagttat
      attttttataaaaatggttcaacaaatgaaagaactaggtgtgaacaaccatctagaaatt
      aataaaaaattaactaatgaagatgcatataaaagagctcgtccatctcacgttacactc
      attgttcaaaatcaagaaggtcttaaaaatttatttaaaatagttagtgcttcattagtt
      aagtattattaccgtacgccaagaattccacgttctcttttaaatgaatatcgagaaggg
 30   atcttgattggtacagcttgtgatgagggtgaattattcacagcagtaatgcagaaggat
      cagtcggaagtagaaaaaaatagcaaagttctatgattttatagaagttcaaccgcctgcg
      ctttatcaagatttaatggatagagaattaatacgagataatgaaacgttaacacaaatt
      tacaagcgattaatagatgctggtaaaagcgctaatatcccagtgattgctactggtaac
      gcgcattatctatatgaacatgatgctatagccagaaaaattttaattgcatcccaacca
 35   gggaatccattaaatcgttcaacattaccagaagctcactttagaaccactgatgaaatg
      ttagatgattttcacttcttaggtgaagaaaaagcatatgaaatcgttgtaacaaataca
      aatgagctcgctaataaaattgaaaaagtggttcctataaaagataaactatttacgcca
      agaatggatggggctaatgaagaaattcgtgagttgagttattctaatgcgaaaaaacta
      tatggtgaagatttaccacaaattgttatagatcgccttgaaaaggaattagatagtatt
 40   attggtaatggcttttctgttatttacctcatatctcaacgtttggtgaagaaatcgcta
      gatgatggttatttagttggatcgcgtggttcagttggttctagtttcgtagcaacaatg
      actgaaattacagaagttaatccgcttccaccacactacatttgttcacattgtaagaca
      agtgagttctttgatgatggtcggttggatctggattcgatttaccagataaaaaatgt
      cctacttgtggtaatgaattaattaaagaaggacaagatatccctttttgagacattcctt
 45   ggattaaaggagataaagttccagatattgatttgaactttagtggtgaatatcaacct
      aacgctcataattacacaaaagtattgtttggtgaagataaagtatttcgtgctggaaca
      ataggtactgttgctgaaaaaacagcttttggtttcgtaaaaggttacttaaatgatcaa
      ggtattcacaaacgtggtgctgaaattgatcggttggttaaaggttgtacaggggtcaaa
      cgtacaactggtcaacatcctggaggaatcattgttgtaccggattacatggatatttat
 50   gattttacaccgattcaattcccagcagacgaccaaagtgcagcgtggatgacaacccat
      ttcgacttccattcaatacacgataatgtcttaaaattagatatattaggacatgatgac
      ccaacgatgattcgtatgttacaagacttatcaggaattgaccccaaaactataccagta
      gatgataaagaaacaatgcaaatatttagtggtcctgagagtttaggtgttacagaagac
      gaaatattatgtaagacaggtacatttggtgtaccagaatttggtactggatttgtacgt
 55   caaatgcttgaagatactaagccaacgacattctcagaattagttcaaatttcaggttta
      tctcatggtacggacgtttggttaggtaatgcacaagagttaattcgtcaagggatatgt
      gacttatctagtgtgataggctgtcgtgatgatatcatggtatatctgatgtatgctgga
      cttgaaccgtcaatggcttttaaaacgatggaatttgtacgtaaaggtcgtggcttaaca
      gatgaaatggttgaagcgatgaaggaaaataacgtgccagattggtatttagattcttgt
```

                                   511

```
      cgtaaaattaaatatatgttccctaaagctcatgccgctgcttatgtactgatggctgta
      agaattgcatactttaaagtacatcatccactatattattatgcagcatactttaccata
      agagcttccgattttgaccttataacaatgattaaagataaaacgagtattcgtaataca
      gttaaagatatgtattcacgatatatggatttagggaaaaaagagaaagatgtattaact
   5  gtattagaaataatgaatgaaatggcgcatcgaggttttcgattgcaaccgattagttta
      gaaaaaagccaagcttttgacttcatcattgaagggggatacattgattcctccattcatt
      tcagtgccaggacttggagaaaacgttgcacaaagaattgttgaagcgagagaagaggga
      ccatttttatccaaagaagatttaaataaaaaagccggcttatctcaaaaggttattgac
      tatttagatgaattaggctcattgccagatttacctgacaaggcacaattgtcgatattt
  10  gatatgtaa
```

Sequence 2052

```
VVIILAMTNREKFKVLADQIKISNQLEQDILEQGELTRIDVSNKNRTWTFQISLPHFLSH
EDYLLFTHAIEEEFKEIATVAIDFSIKDTNNQDEFALKYFGHCIDQTRLSPKVKGQLKQK
KLIMSGNVLKVLVSNDIERNHFDKACNGSLVKAFRQCGFEIDKVVFETDSTNHDDDLASL
EAIIQEDEQSAREATEKLEKMKAEKAKQQDNNESTVEKCQIGKPIQIENIKPIESIIEE
EFKVAILCVIFDINLKELKSGRHIVELKVTDYTDSLVLKMFTRKNKDDLDHFKALSVGKW
VRAQGRIEEDTFVRDLVMMMSDIEEIKKTPKQDKAEDKRVEFHLHTSMSQMDGIPNISAY
VEQAAKWGHQALAVTDHNVVQAFPDAHNAAEKHGIKMIYGMEGMLVDDGVPIAYKPTDRN
LKDATYVVFDVETTGLSNQYDKIIELAAVKVHNGEIIDKFERFSNPHERLSETIINLTHI
TDDMLTDAPEIEEVLTEFKEWVGDAIFVAHNASFDMGFIDTGYERLGFGPSTNGVIDTLE
LSRTINTEYGKHGLNFLAKKYGVELTQHHRAIYDTEATAYIFIKMVQQMKELGVNNHLEI
NKKLTNEDAYKRARPSHVTLIVQNQEGLKNLFKIVSASLVKYYYRTPRIPRSLLNEYREG
ILIGTACDEGELFTAVMQKDQSEVEKIAKFYDFIEVQPPALYQDLMDRELIRDNETLTQI
YKRLIDAGKSANIPVIATGNAHYLYEHDAIARKILIASQPGNPLNRSTLPEAHFRTTDEM
LDDFHFLGEEKAYEIVVTNTNELANKIEKVVPIKDKLFTPRMDGANEEIRELSYSNAKKL
YGEDLPQIVIDRLEKELDSIIGNGFSVIYLISQRLVKKSLDDGYLVGSRGSVGSSFVATM
TEITEVNPLPPHYICSHCKTSEFFDDGSVGSGFDLPDKKCPTCGNELIKEGQDIPFETFL
GFKGDKVPDIDLNFSGEYQPNAHNYTKVLFGEDKVFRAGTIGTVAEKTAFGFVKGYLNDQ
GIHKRGAEIDRLVKGCTGVKRTTGQHPGGIIVVPDYMDIYDFTPIQFPADDQSAAWMTTH
FDFHSIHDNVLKLDILGHDDPTMIRMLQDLSGIDPKTIPVDDKETMQIFSGPESLGVTED
EILCKTGTFGVPEFGTGFVRQMLEDTKPTTFSELVQISGLSHGTDVWLGNAQELIRQGIC
DLSSVIGCRDDIMVYLMYAGLEPSMAFKTMEFVRKGRGLTDEMVEAMKENNVPDWYLDSC
RKIKYMFPKAHAAAYVLMAVRIAYFKVHHPLYYYAAYFTIRASDFDLITMIKDKTSIRNT
VKDMYSRYMDLGKKEKDVLTVLEIMNEMAHRGFRLQPISLEKSQAFDFIIEGDTLIPPFI
SVPGLGENVAQRIVEAREEGPFLSKEDLNKKAGLSQKVIDYLDELGSLPDLPDKAQLSIF
DM*
```

Sequence 2053
Contig_0702_pos_5647_4577,
is similar to (with p-value 0.0e+00)
>sp:sp|P32727|NUSA_BACSU N UTILIZATION SUBSTANCE PROTEIN A H
OMOLOG (NUSA PROTEIN). >pir:pir|C36905|C36905 nusA homolog –
 Bacillus subtilis >gp:gp|Z18631|BSORF1T7A_2 B.subtilis infB
-nusA operon. NID: g49314.

```
atggacgaaggttcatttagagtgattgcacgtaaagaagtcgtagaagaagtgtttgat
gacagagatgaagttgatttaagtactgctttagtcaaaaatcctgcctatgaagtagga
gatatttatgaacaagatgtaacaccgaaagacttcggacgtgtaggagctcaagcagct
aagcaagctgtgatgcaacgacttagagacgcagaaagagaaattttatatgatgaattt
atcgataaagaagaagatattctaacaggtgtgattgaccgtgtagaccatcgctatgta
tatgtgaatttaggaagaattgaagctgtgctgtcagaagctgaaagaagtcctaatgag
aaatatattcctaatgaacgtatcaaggtgtacgtaaataaagttgaacagactacaaaa
ggtccacaaattacgtatcaagaagtcatcctggattactaaaacgcttattcgaacaa
gaagttccagaaatttatgatggtactgttattgttaaatcagtagcgcgtgaagctgga
gatcgttctaaaattagcgtgtattctgataatcctgatatagatgctgttggcgcatgt
gtaggttctaaaggagcacgagtagaagcggttgttgaagaacttggtggcgaaaaatc
gatatcgtccaatgggatgaagatccgaaagtatttgttcgtaatgctttaagtccatca
caagttttagaagtaattgttgatgaagagaatcaatcaactgtagttgtagttcctgat
taccaattatccttagctataggtaaaagagggcaaaacgcacgtttagctgctaaatta
```

acaagttggaagatagatattaaatcagaatctgatgcccgagaagctggaatttatcct
gttattgaatcagaagaagttgcagatgaaattgttaattccggtgacgaagatgttgag
tttgataatgttaacttggaagagacaaacttaactagtacagaattagctgctgaaaat
gatgaagataaaaaagataaaacagaagaagataatgacacagaatcatag

Sequence 2054
MDEGSFRVIARKEVVEEVFDDRDEVDLSTALVKNPAYEVGDIYEQDVTPKDFGRVGAQAA
KQAVMQPLRDAEREILYDEFIDKEEDILTGVIDRVDHRYVYVNLGRIEAVLSEAEASPNE
KYIPNERIKVYVNKVEQTTKGPQIYVSRSHPGLLKRLFEQEVPEIYDGTVIVKSVAREAG
DRSKISVYSDNPDIDAVGACVGSKGARVEAVVEELGGEKIDIVQWDEDPKVFVRNALSPS
QVLEVIVDEENQSTVVVVPDYQLSLAIGKRGQNARLAAKLTSWKIDIKSESDAREAGIYP
VIESEEVADEIVNSGDEDVEFDNVNLEETNLTSTELAAENDEDKKDKTEEDNDTES*

Sequence 2055
Contig_0702_pos_4266_3958,
is similar to (with p-value 3.0e-21)
>sp:sp|P32729|YLXQ_BACSU PROBABLE RIBOSOMAL PROTEIN IN NUSA-
INFB INTERGENIC REGION (ORF4). >pir:pir|E36905|E36905 hypoth
etical protein 2 (infB 5' region) - Bacillus subtilis >gp:gp
|Z18631|BSORF1T7A_4 B.subtilis infB-nusA operon. NID: g49314
. >gp:gp|Z99112|BSUB0009_132 Bacillus subtilis complete geno
me (section 9 of 21): from 1598421 to 1807200. NID: g2633902

atgaaattttaatttgcttggtttagctatgagagctggtaaaatcaaaagtggcgaa
tcggtcatcttaaatgagcttaaaaagaatcaaataaaacttgtcatattagctagcgat
gcatctagtaacactctaaaacaaatgaataataaatgtaatagttaccaagtgccatta
aaagtgtttggtactagaaatgaattaggggttagcaataggtaaaagcgatagattaat
attggtataacagataatggttttgcaaaaaaattgttatcaatgatagatgaatatggt
aaggagtga

Sequence 2056
MKIFNLLGLAMRAGKIKSGESVILNELKKNQIKLVILASDASSNTLKQMNNKCNSYQVPL
KVFGTRNELGLAIGKSDRVNIGITDNGFAKKLLSMIDEYGKE*

Sequence 2057
Contig_0702_pos_3953_1791,
is similar to (with p-value 0.0e+00)
>sp:sp|P17889|IF2_BACSU TRANSLATION INITIATION FACTOR IF-2.
>pir:pir|A35269|A35269 translation initiation factor IF-2 -
Bacillus subtilis >gp:gp|M34836|BACPSIF2A_1 B.subtilis prote
in synthesis initiation factor 2 (infB) gene, complete cds.
NID: g143358. >gp:gp|Z18631|BSORF1T7A_5 B.subtilis infB-nusA
operon. NID: g49314.
atgagtaaaaaaagaatttacgaatatgcgaaagaattaaatctaaagagtaagagatt
atagatgagttaaaaagtatgaatgttgaagtgtcaaatcatatgcaagctttagaagaa
gaacaaatcaaagcattagataaaaaatttaaagcctctcaagcgaaagacactaataaa
caaatactcaaaataatcaccaaaaatctaataataaacaaaattctaacgataaagaa
aaacaacaaagtaagaataatagtaaaccaacgaagaaaaagaacaaaacaacaaagga
aaacagcaaaataaaaacaataaaactaataagaatcaaaaaaacaataaaaataaaaag
aataataaaataataaacctcaaaatgaggtagcagaaacaaaagaaatgccctctaaa
atcacttatcaagaaggcataactgtcggtgagttagctgaaaagctaaatgtagaatca
gctggtattattaaaaaattgttcttactaggtattatggctaatatcaatcaatcattg
gatgaagaaacattagaattaattgcagatgactatggcgttgaaatagagaagaagta
gtcgttgatgaagaagatttatcaatttattttgatgatgagactgatgattctgatgca
attgaacgtccagcagttgttacaatcatgggccacgtagaccatggtaaaacgacttta
ttagattctattcgtaacactaaagttacagaaggagaagctggcggaatcactcaacat
attggtgcttatcaaattgaaaattcaggtaaaaaaattacgttcttagatactcctgga
catgctgcatttacgactatgcgtgcacgtggtgctcaagttactgatattacaatttta
gtcgtggccgctgatgatggtgtgatgcctcaaacaattgaagctataaatcacgctaaa

gaagcagaagtacctacgattgttgcagtaaacaaaattgataaaccaactgctaaccct
gatcgtgttatgcaagaactcactgagtatggattaattccagaagactgggcggtgac
acaatctttgtaccactatctgcattgagtggagacggtattgatgatttattagaaatg
atcggtttagtagcggaggtacaagaacttaaagctaatcctaataaacaagctgtaggt
actgtgattgaggctgaattagataaatcacgaggtccagctgcatctttacttgttcaa
aatggtactttaaacgttggagatgcaattgttgtaggtaatacttatggacgtatacgt
gcaatggttaatgatttaggaaaaagaattaaatctgccggtccttcaacacctgtagaa
attactggtattaacgatgttccacttgcaggtgatcgttttgttgtatttggtgatgaa
aaacaagcacgtcgaattggtgaagcacgtcatgaggcaagtgtcatacagcaacgtcaa
gaaagtaaaaatgtttcattagacaatttatttgagcaaatgaaacaaggtgaaatgaaa
gatttaaatgtcatcattaaaggtgatgtacaaggttcagttgaagcattggccgcatct
ctaatgaaaatagatgttgaaggtgtgaatgtacgaattattcatacagctgttggtgct
atcaatgaatcagatgttacattagcaaatgcatcaaatggtattattattggttttaat
gtacgcccagatgcaggtgcgaaacgtgcggctgaagctgaaatgtagatatgcgatta
cacagagttatctataatgttattgaagagatagaatcagctatgaaaggtttacttgac
ccagaatttgaagagcaagtcattggacaagctgaagtgcgtcaaacatttaaagtttct
aaagttggtacaattgctggtagttatgtgactgaaggtaaaatcactcgtaacgctggt
gtacgcgtaattagagatggtatcgtgttatttgaaggtgaacttgacacattaaaacgt
ttcaaagatgatgctaaagaagtagctcaaggctatgaatgtggtattacaattgaaaaa
tataatgatctcaaagaaggagacattattgaagcgtttgaaatggtagaaattcaaaga
taa

Sequence 2058
MSKKRIYEYAKELNLKSKEIIDELKSMNVEVSNHMQALEEEQIKALDKKFKASQAKDTNK
QNTQNNHQKSNNKQNSNDKEKQQSKNNSKPTKKKEQNNKGKQQNKNNKTNKNQKNNKNKK
NNKNNKPQNEVAETKEMPSKITYQEGITVGELAEKLNVESAGIIKKLFLLGIMANINQSL
DEETLELIADDYGVEIEKEVVVDEEDLSIYFDDETDDSDAIERPAVVTIMGHVDHGKTTL
LDSIRNTKVTEGEAGGITQHIGAYQIENSGKKITFLDTPGHAAFTTMRARGAQVTDITIL
VVAADDGVMPQTIEAINHAKEAEVPTIVAVNKIDKPTANPDRVMQELTEYGLIPEDWGGD
TIFVPLSALSGDGIDDLLEMIGLVAEVQELKANPNKQAVGTVIEAELDKSRGPAASLLVQ
NGTLNVGDAIVVGNTYGRIRAMVNDLGKRIKSAGPSTPVEITGINDVPLAGDRFVVFGDE
KQARRIGEARHEASVIQQRQESKNVSLDNLFEQMKQGEMKDLNVIIKGDVQGSVEALAAS
LMKIDVEGVNVRIIHTAVGAINESDVTLANASNGIIIGFNVRPDAGAKRAAEAENVDMRL
HRVIYNVIEEIESAMKGLLDPEFEEQVIGQAEVRQTFKVSKVGTIAGSYVTEGKITRNAG
VRVIRDGIVLFEGELDTLKRFKDDAKEVAQGYECGITIEKYNDLKEGDIIEAFEMVEIQR
*

Sequence 2059
Contig_0702_pos_1521_1138,
is similar to (with p-value 2.0e-34)
>sp:sp|P32731|RBFA_BACSU RIBOSOME-BINDING FACTOR A (P15B PRO
TEIN). >pir:pir|G36905|G36905 protein P15B homolog - Bacillu
s subtilis >gp:gp|Z18631|BSORF1T7A_7 B.subtilis infB-nusA op
eron. NID: g49314. >gp:gp|Z99112|BSUB0009_135 Bacillus subti
lis complete genome (section 9 of 21): from 1598421 to 18072
00. NID: g2633902.
gtgagtaaagtaaaaataaagagaggtgagatgatgaataatataagagcagaacgtgta
ggagaacaaatgaaacaggaaatcatggacattgttaataataaagttaaagaccctaga
gttggttttttaacaattactgatgttgaactaaccaatgacctttcacaagcaaaggta
tatttaacagtgttagggaatgataaagaagttgataatacgtttaaagctttgcataaa
gcaactgggtttataaaatctgaacttggttctcgaatgcgcctaagaattatacctgag
ttaacattcgaatatgatgaatctatcgaatacggtaataagatagaacgcatgattcaa
gagttacacaaaaatgataaataa

Sequence 2060
VSKVKIKRGEMMNNIRAERVGEQMKQEIMDIVNNKVKDPRVGFLTITDVELTNDLSQAKV
YLTVLGNDKEVDNTFKALHKATGFIKSELGSRMRLRIIPELTFEYDESIEYGNKIERMIQ
ELIKNDK*

Sequence 2061
Contig_0702_pos_1000_83,
is similar to (with p-value 6.0e-46)
>gp:gp|X92946|LLLPK214_14 Lactococcus lactis sp. lactis plas
5   mid pK214, complete sequence. NID: g2467210.
atgtataatggcatactaccggtatttaagaaacgaggtttaacaagtcacgacgtcgtt
tttaaattacgtaaaattttaaaaatgaaaaaaattggtcatacaggaacattagatcct
gaagttaatggtgtgttaccaatttgtttaggcgatgcgacaaaagtcagtgattatatc
atggaaatgggaaaaacttatcacgctatgataacgctaggaaagagtacgactactgaa
10 gaccaaactggagatattttagaaactagggctgttgataagaatgatattaatgaagat
acgattgaccaagtgttgcagcaatttgaggggcatattcaacaaattccgcctatgtat
tcttctgttaaagtaaatggaagaaaattatatgaatatgcgagaaataatgaaactgtc
gaacgccctaaacgacaagttttattaaagatatacatagaatatctgaagttactttt
caggagcagacatgtcattttgaagttgaagtaacatgtggtaaaggaacttatattaga
15 actttagctacagatattggacttaaacttggttttccagctcatatgtcacgtctaact
agaattgcttctggcggttttcaattagaaagtagtttaacgattgatcaaattaaagaa
ttacatgagcatgattcattacataatgaattgtttcctatagaatatggcttaaaaggt
ctgaaatcattccaagtgaaagattcaaatttaaaaagaaatctgtaacggtcaaaaa
tttcataaaaaagtgttaagtcaaaatgttaaagaaccttttatatttgtcgatagtagc
20 actcaaaaagttttagcaatatatatagttcatccagataaaccttatgaaataaaacct
aaaaaagtttttaattaa

Sequence 2062
MYNGILPVFKKRGLTSHDVVFKLRKILKMKKIGHTGTLDPEVNGVLPICLGDATKVSDYI
25 MEMGKTYHAMITLGKSTTTEDQTGDILETRAVDKNDINEDTIDQVLQQFEGHIQQIPPMY
SSVKVNGRKLYEYARNNETVERPKRQVFIKDIHRISEVTFQEQTCHFEVEVTCGKGTYIR
TLATDIGLKLGFPAHMSRLTRIASGGFQLESSLTIDQIKELHEHDSLHNELFPIEYGLKG
LKSFQVKDSNFKKKICNGQKFHKKVLSQNVKEPFIFVDSSTQKVLAIYIVHPDKPYEIKP
KKVFN*
30

Sequence 2063
Contig_0703_pos_784_1236,
putative peptide of unknown function
atgtatgatgaattagaggttaataaaagtcgtttaaaaactgtaattttaatgaaggt
35 atttttaagaatatagaagcaatttgtaattgtaaatttacaacgtgcgggtttaataat
tgtattttcgaagatgttcattttacaaaaaccaatttaaagattcaacatttgtgaat
acaccatttgatcaatccgtatttaatagtactttattccaaaatgcaatgttcgatagc
aatctcattcgtagcgtaaaatggactgatatcattttttaaaaacgtttctttcaaaaat
gtagaaattgaaggaacaacatttaaagatgtaaaattcaaaaattgtgagttcaaaat
40 gtaattattactaattcaactatgtcgcaaaagttaatgaatgaattacaaaaacaagat
gttactttagaaaatatagacacttctatttaa

Sequence 2064
MYDELEVNKSRFKNCNFNEGIFKNIEAICNCKFTTCGFNNCIFEDVHFYKNQFKDSTFVN
45 TPFDQSVFNSTLFQNAMFDSNLIRSVKWTDIIFKNVSFKNVEIEGTTFKDVKFKNCEFKN
VIITNSTMSQKLMNELQKQDVTLENIDTSI*

Sequence 2065
Contig_0703_pos_4474_5274,
50 putative peptide of unknown function
atgaaagatagggttgaaacagaagaatatgctagaaatcaattaatctctaaaaattca
attttaagtgaagaaaattatcattgaaaaaccaaatgttaagtacaaacaatgacgtc
ggtcaacacgcttttaaaaacgccaagcgtgaattaagaaaaatattaaatagatttaaa
gaagagggtcgtttacgatcatatacaattgttcctacgagtaatttggctgttaaacat
55 cccctttcgaatatgcacgttcattcgatttattatcattactgatgttggtttgata
aatgtggatgttaaaaattggaaccaaaaaacgttttatcattttgatgtgccagatcaa
catcttgaagaggacaaccacaatataataccgaaaaagttgtcggtcattatattagc
aatcgatatcatagtcagtttaaaacaacacgttctggtgtctatactttttattgagatt
ttacaggataatcgtgtaatatatgaattttatgaccacgatccatacgataaagccgca

```
aacaatgcaaaagcattaaaagataaaattgaaaatgattataattttaaaattcaaagt
attggcgtcatatattttagtgatggtagcgttaatattattgaaggatccgacgagagt
gataaatacgtcgacaccgtatctacaccgatatcacttgaaaaagtaattgaggaagct
atcgatttatctaagcacccccttactgataaacaaatcgaagaaatttctgaaaacttt
```
5   `aaacaacatatgaataattaa`

Sequence 2066
MKDRVETEEYARNQLISKNSILSEENLSLKNQMLSTNNDVGQHAFKNAKRELRKILNRFK
EEGRLRSYTIVPTSNLAVKHPLFEYARSFDFIIITDVGLINVDVKNWNQKTFYHFDVPDQ

10   HLEEGQPQYNTEKVVGHYISNRYHSQFKTTRSGVYTFIEILQDNRVIYEFYDHDPYDKAA
NNAKALKDKIENDYNFKIQSIGVIYFSDGSVNIIEGSDESDKYVDTVSTPISLEKVIEEA
IDLSKHPLTDKQIEEISENFKQHMNN*

Sequence 2067
15   Contig_0703_pos_3987_3481,
putative peptide of unknown function
```
atgcatctatttctgaatgatgattttatattaccatcgaaaatagatatttgtgttgag
cgcaaaaagttattggaaattatcagggtgattcctgaagaatttacgattcactattat
gataatcagttatatgaaagattaagagaatctttatcttatcagatgttgaacatgtt
```
20   ```
aaagtctttaaaaacgatacagaagtcatgacaatatttgtttatgatgttgtaaatgat
gaatggctatttaggttagatcatcatatacgtttaccaaaaaataataatatatttcat
tctttaagttggaacgtagattatattaagccagaaatagttcttatgtatgattiaatg
agtgaacaaaaatatcatcagtttagtaattataaagctgttattgattctcttagttat
tatcaattctatattttaaaattggtagtaggtgaacagcgtattaaaaaagctatagta
```
25     `aatagttccactaaaaagatatcttaa`

Sequence 2068
MHLFLNDDFILPSKIDICVERKKLLEIIRVIPEEFTIHYYDNQLYERLRESLSLSDVEHV
KVFKNDTEVMTIFVYDVVNDEWLFRLDHHIRLPKNNIYFHSLSWNVDYIKPEIVLMYDLM

30     SEQKYHQFSNYKAVIDSLSYYQFYILKLVVGEQRIKKAIVNSSTKKIS*

Sequence 2069
Contig_0703_pos_2163_1420,
putative peptide of unknown function

35     ```
atgtcgtcattatgccttttgcctggagctgaaacgataatgattttacgttcagggtct
tcattcactataggacaatcgttttttacgctgtccctaggtacaaccggaatgattact
tatgcaagctatgcacctaaaaatatgacgataaagtcttcagcacttttcaattgtcgta
atgaatattttaatttctgtcttggctggattagctatatttcctgcgcttaaaacattt
ggttaccaaccccaagaaggccctggcttattatttaaggttttaccactagtatttagc
```
40   ```
gaaatgacttttggtacattcttttactttatattttttactattattcttatttgcggca
ttaacgtcttctatatcattattagagttaaatgtatctaattttactaaaaatgataat
agtaaaagacaaaaagtggcaatcataggtagtatacttgtatttatcattagtatccca
gcaacattatctttttagtagtctaagtcatttgcgttttggcgctggtacgatatttgat
aatatggattttattgtatctaatattcttatgccattaggggcactaggaacaacatta
```
45     ```
gtggttggccaattactagataaaaaattattaaaagaaagctttgggaaagacaaattc
aacctatttttaccgtggtattatttaattaagttcatcatgcctattgttattatttta
gtatttatagttcaattattttaa
```

Sequence 2070
50   MSSLCLLPGAETIMILRSGSSFTIGQSFFTLSLGTTGMITYASYAPKNMTIKSSALSIVV
MNILISVLAGLAIFPALKTFGYQPQEGPGLLFKVLPLVFSEMTFGTFFYFIFLLLFLFAA
LTSSISLLELNVSNFTKNDNSKRQKVAIIGSILVFIISIPATLSFSSLSHLRFGAGTIFD
NMDFIVSNILMPLGALGTTLVVGQLLDKKLLKESFGKDKFNLFLPWYYLIKFIMPIVIIL
VFIVQLF*

55
Sequence 2071
Contig_0703_pos_0_399,
is similar to (with p-value 9.0e-32)
>gp:gp|U93874|BSU93874_1 Bacillus subtilis cysteine synthase
```

(yrhA), cystathionine gamma-lyase (yrhB), YrhC (yrhC), YrhD
(yrhD), formate dehydrogenase chain A (yrhE), YrhF (yrhF),
formate dehydrogenase (yrhG), YrhH (yrhH), regulatory protei
n (yrhI), cytochrome P450 102 (yrhJ), YrhK (yrhK), hypotheti
5 cal protein YrhL (yrhL), putative anti-SigV factor (yrhM), R
NA polymerase sigma factor SigV (sigV) and YrhO (yrhO) genes
, complete cds, and YrhP (yrhP) gene, partial cds. NID: g193
4604. >gp:gp|Z99117|BSUB0014_206 Bacillus subtilis complete
genome (section 14 of 21): from 2599451 to 2812870. NID: g26
10 34966.
atgattgcatacgatttgataggacaaactccattagttttattagaaagctttagtgac
gagaatgttaaaatatacgccaaacttgagcaatttaatcctggtggtagcatcaaagac
cgtctagggaagtacttaattgaaaaagcaatagatgaaggacgacttaaagaaggggat
acaatagtcgaagcgactgctggtaatacaggcattggacttgctattgcttctaatcgg
15 cacaaagtaaaatgtatcatctttgctccagaaggatttgcagaagaaaaatttcaatt
atgaaagcattgggtgcagatgttagacgtaccccaaagctgagggaatgactggcgca
cagcaagaggcgttggcatacgcaacacgatatgTACTA

Sequence 2072
20 MIAYDLIGQTPLVLLESFSDENVKIYAKLEQFNPGGSIKDRLGKYLIEKAIDEGRLKEGD
TIVEATAGNTGIGLAIASNRHKVKCIIFAPEGFAEEKISIMKALGADVRRTPKAEGMTGA
QQEALAYATRYVL

Sequence 2073
25 Contig_0704_pos_804_1115,
putative peptide of unknown function
gtgttatacagaaaaatggttgctaaaataaaacttaaacgcctatacaaaccctattca
caaactttattttttgttgttcattttaaaatatttttactattttactataatccttgtt
ttttattccaaatactttacatcatccttgggtaaagaaggattgattcttatcctcatt
30 aataataaatgtaattataaaagcctttccgtgaactcaataagtctgaattctaaaaag
cgaaacagaaatcttatcatattcttctttgtttcaatccattaccgaaccccaacttgc
tttgtctgttga

Sequence 2074
35 VLYRKMVAKIKLKRLYKPYSQTLFLLFILKYFYYFTIILVFYSKYFTSSLGKEGLILILI
NNKCNYKSLSVNSISLNSKKRNRNLIIFFFVSIHYRTPTCFVC*

Sequence 2075
Contig_0704_pos_2016_0,
40 is similar to (with p-value 3.0e-76)
>sp:sp|P47994|SECA_STACA PREPROTEIN TRANSLOCASE SECA SUBUNIT
. >pir:pir|S47149|S47149 secA protein - Staphylococcus carno
sus >gp:gp|X79725|SCSECA_2 S.carnosus (TM300) secA gene. NID
: g499333.
45 atgacctgtttagatagaataactggtcgtatgctacctggaacaaagcttcagtctggt
ttacatcaagctatagaggctctggaaaatgttgaaatttctcaagatatgagtgtgatg
gcaaccataacattccaaaacttatttaagcaatttgatgaattttcaggtatgactgga
acaggtaaattaggggaaaaagaattctttgatttatattcaaaagttgttatagagatt
ccgactcacagtccgattgaacgagatgatagacctgatagagtatttgctaatggtgac
50 aaaaagaacgatgcaattttaaagacagtgattggtatacatgaaactcaacaacctgtg
ttactaattacacgtactgcagaagcggcagaatattttcagctgagttatttaaacgt
gatatacccaacaatttattaatcgctcaaaatgtagctaaagaggcacaaatgattgct
gaggcgccacaattatctgcagttactgttgctacaagtatggcagggcgtggaactgat
ataaagttatcaaaagaggttcatgatatcggtggcttagcagtgattattaatgaacat
55 atggataatagccgtgttgatcgtcaattaagaggacgctcaggtcgccaaggagatcct
ggatattcacagattttttgtatcacttgatgatgatttagtaaaacgttggagtaactct
aacttggcagaaaataaaaaacctccaaacgatggatgcatctaaactagaaagtagtgca
ctctttaaaaaacgtgtaaagtcaattgttaataaagcgcaacgtgtatctgaagagact
gctatgaaaaatagagaaatggcaaatgaattcgaaaaaagtattagtgttcaacgagat

517

```
aaaatttatgctgaacgtaatcacatacttgaagcaagcgattttgatgattttaatttt
gaacagcttgcacgagatgtgtttacaaaagacgttaaaaatcttgacttaagtagtgaa
cgtgcacttgtgaattatatatacgaaaacttaagttttgtcttcgatgaagatgtatca
aatattaatatgcaaaatgatgaagaaatcatacaattcttaatacaacaatttacccaa
```
5
```
caatttaacaatcgtttagaagttgctgctgattcatatttaaaacttcgtttcattcaa
aaatcaattttgaaagcgatagatagcgaatggattgaacaagtagataacttacaacaa
cttaaagccagtgtaaacaatcgacaaaatggacagcgtaatgtcatttttgaatatcat
aaagtggctcttgaaacgtatgaatatatgtctgaagatataaaaaggaagatggttaga
aatttatgtttaagtattctagcctttgataaggacggagata
```
10

Sequence 2076
```
MTCLDRITGRMLPGTKLQSGLHQAIEALENVEISQDMSVMATITFQNLFKQFDEFSGMTG
TGKLGEKEFFDLYSKVVIEIPTHSPIERDDRPDRVFANGDKKNDAILKTVIGIHETQQPV
LLITRTAEAAEYFSAELFKRDIPNNLLIAQNVAKEAQMIAEAGQLSAVTVATSMAGRGTD
```
15
```
IKLSKEVHDIGGLAVIINEHMDNSRVDRQLRGRSGRQGDPGYSQIFVSLDDDLVKRWSNS
NLAENKNLQTMDASKLESSALFKKRVKSIVNKAQRVSEETAMKNREMANEFEKSISVQRD
KIYAERNHILEASDFDDFNFEQLARDVFTKDVKNLDLSSERALVNYIYENLSFVFDEDVS
NINMQNDEEIIQFLIQQFTQQFNNRLEVAADSYLKLRFIQKSILKAIDSEWIEQVDNLQQ
LKASVNNRQNGQRNVIFEYHKVALETYEYMSEDIKRKMVRNLCLSILAFDKDGDX
```
20

Sequence 2077
```
Contig_0704_pos_1840_1244,
is similar to (with p-value 3.0e-87)
>sp:sp|P24277|RECR_BACSU RECOMBINATION PROTEIN RECR. >gp:gp|
```
25
```
D26185|BAC180K_85 B. subtilis DNA, 180 kilobase region of re
plication origin. NID: g467326. >gp:gp|X17014|BSRECM_3 Bacil
lus subtilis dnaZX and recR genes and two unidentified readi
ng frames. NID: g453238. >gp:gp|Z99104|BSUB0001_21 Bacillus
subtilis complete genome (section 1 of 21): from 1 to 213080
```
30
```
. NID: g2632267.
atgcattatccagaacctatatcaaagcttatcgatagtttttatgaaactgccaggcatt
ggaccaaagacggctcaacgtctggcttttcatactttagatatgaaagaagacgatgtt
gttaagtttgctaaagcactagttgatgttaaaagagaacttacctattgtagtgtttgt
gggcatattacagaaaatgatccttgttatatatgtgaagataaacagcgagatcgttct
```
35
```
gtcatatgtgtagttgaagatgacaaggatgtcatagcaatggaaaaaatgcgtgaatat
aaaggtttatatcacgtgcttcatggttcgatttcaccaatggatggtattgggcctgaa
gacatcaatatacctgcattagttgaacgcctcaaaaacgatgaggtgaaagagcttata
ttagctatgaatcctaacctagaaggcgagtctactgcaatgtatatatctaggttggtt
aaaccaattgggattaaagtcacaagactggcacaaggtttatctgtaggcggcgattta
```
40
```
gaatatgctgatgaagtgactttatctaaagcaattgcaggtagaacggaaatgtaa
```

Sequence 2078
```
MHYPEPISKLIDSFMKLPGIGPKTAQRLAFHTLDMKEDDVVKFAKALVDVKRELTYCSVC
GHITENDPCYICEDKQRDRSVICVVEDDKDVIAMEKMREYKGLYHVLHGSISPMDGIGPE
```
45
```
DINIPALVERLKNDEVKELILAMNPNLEGESTAMYISRLVKPIGIKVTRLAQGLSVGGDL
EYADEVTLSKAIAGRTEM*
```

Sequence 2079
```
Contig_0706_pos_183_1418,
```
50
```
is similar to (with p-value 0.0e+00)
>sp:sp|P55189|YBAR_BACSU HYPOTHETICAL 46.4 KD PROTEIN IN RRN
G-FEUC INTERGENIC REGION. >gp:gp|D84213|BACTHRTRNA_3 Bacillu
s subtilis genome, trnI-feuABC region. NID: g1256147.
gtggcgcttttagtcacacctttagttaaagaacatggtgttgaatatcttttttgctgct
```
55
```
acgatattgatggggttaatacaattacttttaggaatacttaaagtcggtcgtttaatg
aaatttattccccgtccagttatgattggatttgttaatgcattgggtattatgatttttt
atgtctcaaatagagcatattttttaatatatctatagcaacttatatatacgttattata
actttactaattgtatatgttattcctagattttataaagctatacctgccccattaata
gctataattgtttttaacagcattgtatatgtatacaggatctgacgtaagaactgtaggt
```

```
gatttaggtaatattaagcaaactttgcctcatttcttaatccctaacattccttttaat
ttagaaacacttcaaattattttcccatattcattatcaatggctattgtgggtctagta
gaaagcctacttacagctaaaattgtagatgatgcaactgatacacacagtagtaaaaac
aaagaatcaagaggtcaagggattgctaatattattactggtttctttggcggtatggga
ggttgtgccatgattggacaatcagtaatcaatgttaagtcaggtgcaaatagtagattg
tctacatttacagcaggtattgtacttatgtttatgattattgttctcggtggtgtagtc
gttcaaatcccgatgccaattcttgcagggataatggttatggtttctatcgggactgtt
gattggaattcatttaagtatataaaaaaagcacctaaaacagatgcattcgttatggtt
ttaacggttattattgtattaataacacataacttagctatcggtgtagttgtaggtgtt
gttttcagtgctatattctttgcgactaaaatttctaaagttgaagttatttataaagag
ttaggtaagcagcatcgttttttcttttaaaggtcagatattctttgtatcaattgattct
atgatggagaaaattgattttaatattgaagatagtgtcatagtgttgaattttgatcac
gctcacctatgggatgattcagcagtaaatgccattgatacgattgttaggaaatttgaa
gagaaaaacaatattgtgtatgttgaaaaattaaacgcagatagccgaaaaattatttca
gaactaagttatttaaatgaaactcatttaaattaa
```

Sequence 2080
```
VALLVTPLVKEHGVEYLFAATILMGLIQLLLGILKVGRLMKFIPRPVMIGFVNALGIMIF
MSQIEHIFNISIATYIYVIITLLIVYVIPRFYKAIPAPLIAIIVLTALYMYTGSDVRTVG
DLGNIKQTLPHFLIPNIPFNLETLQIIFPYSLSMAIVGLVESLLTAKIVDDATDTHSSKN
KESRGQGIANIITGFFGGMGGCAMIGQSVINVKSGANSRLSTFTAGIVLMFMIIVLGGVV
VQIPMPILAGIMVMVSIGTVDWNSFKYIKKAPKTDAFVMVLTVIIVLITHNLAIGVVVGV
VFSAIFFATKISKVEVIYKELGKQHRFSFKGQIFFVSIDSMMEKIDFNIEDSVIVLNFDH
AHLWDDSAVNAIDTIVRKFEEKNNIVYVEKLNADSRKIISELSYLNETHLN*
```

Sequence 2081
Contig_0706_pos_1422_1844,
putative peptide of unknown function
```
gtgtttttttatgtataattcaattttgcttgcagcagatggttcaaaaaatagtatacgt
gcagcccaagagttactaaattttataggtgattatacaattgtcacaatacttacggtt
gtagatattgaagaatcaaaaacagatgttttacatgatcaccaaggaactaatctaact
caaaaaagggaaagtaaattacaatctattaaagatttatttacagaacacaatgtgaat
tacaaaattaaaattgttcacggtactccaacagataaagtggttgaagtttcaaatagt
ggtgaatatcaagctatcatttttaggtacacgtggcttaaacagttttcaaaaaatggtt
cttggtagtgtaagtcataaagttgctaaacgttctcaaatacctgtaattattgttaaa
taa
```

Sequence 2082
```
VFFMYNSILLAADGSKNSIRAAQELLNFIGDYTIVTILTVVDIEESKTDVLHDHQGTNLT
QKRESKLQSIKDLFTEHNVNYKIKIVHGTPTDKVVEVSNSGEYQAIILGTRGLNSFQKMV
LGSVSHKVAKRSQIPVIIVK*
```

Sequence 2083
Contig_0706_pos_5297_6436,
putative peptide of unknown function
```
atgaaaaaaatatggattttaactataggtatgtttgccttaggtatggatgcttatata
gtagcaggattaataccttcaataagtaaaagttttaataaaagtagctctgctattggg
caaggagtaacagtttttacattgttttttctctatctctgcccccatttttcaacaata
ttagctaaatccccagttaaaaaaatactaataatagcattcagtatatttactttagcc
aatattataaccgcaatatctatgaactacatgctatatatcgtatcaagagcaatcgct
ggtttaggagctggcgtattctcaccaattgcaataagtgcaagcaatcatttagtctcc
gaaaagcataaaggaaaagcaatcgcttttacagtaggcggaatgagtgtaggaactgtt
ataggagttcctctcggactagaaattgccaacatttctaattggcgatttgcaatgttg
gttattattgtcattagtttttattgcattaataagcatatctatattgatgcctaaattt
aaaataggaagctcctccaaatttaaaagatcgttttcaattatttttaaacaagcaagta
ctaagagttatttcggttacattatgcgctgccattgctagtttaggttgtatacttat
ttagccgatattattaaaacaaatacagatacaaaaaatttaactcattaccttacagcg
tggggaataggcggattaataggaagttttggtataggatttattatagatagatttaaa
aatacaagatttgttatgctaattattttaattttactagcattaagttttggtttaatt
```

cctatttctattaacttgcctatattaggtttaattccctttattttatggggagctatg
ggatgggctacacaagctcctcaacaacatatattattgaaaaaacatcctgaatatgga
ggctctgctgtcgctttaaatagttctattaattatttaggcagtgctatgggatcagca
atcggaggaattattttatttaatgctaatagtacaaatgtactaatatatagtgcttta
ggaattactattattggtattttattacaattactaaatttatccctagaaaaaaattaa

Sequence 2084
MKKIWILTIGMFALGMDAYIVAGLIPSISKSFNKSSSAIGQGVTVFTLFFSISAPIFSTI
LAKSPVKKILIIAFSIFTLANIITAISMNYMLYIVSRAIAGLGAGVFSPIAISASNHLVS
EKHKGKAIAFTVGGMSVGTVIGVPLGLEIANISNWRFAMLVIIVISFIALISISILMPKF
KIEAPPNLKDRFQLFLNKHVLRVISVTLCAAIASLGLYTYLADIIKTNTDTKNLTHYLTA
WGIGGLIGSFGIGFIIDRFKNTRFVMLIILILLALSFGLIPISINLPILGLIPFILWGAM
GWATQAPQQHILLKKHPEYGGSAVALNSSINYLGSAMGSAIGGIILFNANSTNVLIYSAL
GITIIGILLQLLNLSLEKN*

Sequence 2085
Contig_0706_pos_5156_4323,
putative peptide of unknown function
atgggggattttatgaattttaattatatgaaggacttcattaaagttgtcgagtacaat
agcttaaataaagcttctagagaattgaatataagtacaccagccttaagtaaaagaatt
agaagtattgaagattattttgattgtcagttatttatagaacttctaagggtattttt
ttaactcaaaaaggaaatgctgtttatcattcgtttttaaaaataaacgagcaatttgaa
gaattaaaaagtaaaataagtgagtctaaagataaaagaattaaattaggagtcatacct
agcttttctttatataagctacatgaaaagaattatattaatgaaagtgtcgttttagtt
attgaaaatagtacttcaattttacttgaggaaatttataaagataataatagatgttgtt
attjgtgatattacacattaaaaaataacctcgctctatacacaagaaatttacacttgag
gactttattgtgtggcttatggtgatgaaaataagtttaaaaacactaaccaagtaagtata
aaatctttaaaaaatgagaatatttatatacaaacaccgccatgtgatacatatgatctt
cttaaaaattattccttgaaagataagttacatattgattatgttgattattatgaaact
atattagctaatgtcaaagcaaataaaggtattactttgactcctcaatctttgactaca
agacttgaaagtatgaacttatatcaaaaaaaattacaaagttataaaagagttgtaggc
gtagtatcacgtgataaagataagatgaataaaattattgatataattcagtaa

Sequence 2086
MGDFMNFNYMKDFIKVVEYNSLNKASRELNISTPALSKRIRSIEDYFDCQLFYRTSKGIF
LTQKGNAVYHSFLKINEQFEELKSKISESKDKRIKLGVIPSFSLYKLHEKNYINESVVLV
IENSTSILLEEIYKDNIDVVIGDITHLKNNSLYTQEIYTEDFIVAYGDENKFKNTNQVSI
KSLKNENIYIQTPPCDTYDLLKNYSLKDKLHIDYVDYYETILANVKANKGITLTPQSLTT
RLESMNLYQKKLQSYKRVVGVVSRDKDKMNKIIDIIQ*

Sequence 2087
Contig_0706_pos_3655_3065,
putative peptide of unknown function
gtgggtgaattcaaattccctagtggatcgataaaagtaacacgtaaaggtattgaaaaa
gatcaggaagagtatttaagacagtttgaaaaagaaatgataaaagtatagaatttgat
gctgatgagatgagcgctaaaatcggagaattatttggagtagaatataaagatggtttg
cctattgataattctggtggtggtggtgctccacctaaatcagaaagttataaccccctca
gtttttaacgctcatctattttctatagtggctagtgaagattctttaaaagaagagtat
aaattttttcggtttcaaagagaccattatctctttatatactgcctatttatctaattat
aaaaaagataagtgggttactaaaggtggtgctccttatcatgtttcaaatagagaagag
agcttagagaagaaagcatttgatattataactgggtattatttagaagctatgaatact
cgtaaagaatggaactttaatactataaaatcttatcaaaagttaataagagttgtgttg
gtttcttatcaagttatagaacgagaacttatggaggaaaatgaagaataa

Sequence 2088
VGEFKFPSGSIKVTRKGIEKDQEEYLRQFEKENDKSIEFDADEMSAKIGELFGVEYKDGL
PIDNSGGGGAPPKSESYNPSVFNAHLFSIVASEDSLKEEYKFFGFKETIISLYTAYLSNY
KKDKWVTKGGAPYHVSNREESLEKKAFDIITGYYLEAMNTRKEWNFNTIKSYQKLIRVVL

VSYQVIERELMEENEE*

Sequence 2089
Contig_0706_pos_2458_2060,
putative peptide of unknown function
atgcaaatgttacttaaagataaagttgacagtgaaggtaaatctattatctatgaagga
caaagaggatctggggggaggagaattagatactcatgatcccatgtatttttcgaattta
attactagtaaaatatttttaattaatcaagtagaaaaacataataagttttttatagct
gaaaagctttttgaaggttatgaaagtaaatgtatggaaaaatatggaacaaattatcac
ggagcaataatagagacaagagatgaaaataggatattagccgatgaaacaatacgtaaa
ttttacttttttaacgaaagacatggaaaaagagaatgaatatttaagcttagatgattta
aataaatttagactttttaatatttgcatttactcattga

Sequence 2090
MQMLLKDKVDSEGKSIIYEGQRGSGGGELDTHDPMYFSNLITSKIFLINQVEKHNKFFIA
EKLFEGYESKCMEKYGTNYHGAIIETRDENRILADETIRKFYFLTKDMEKENEYLSLDDL
NKFRLLIFAFTH*

Sequence 2091
Contig_0707_pos_498_965,
putative peptide of unknown function
atgcaattagatgttttttcctgaattaaacgtagatcaactatctcaaaaagtgagaaaa
atacttaatgcagaaccagataaatatattaaaaatagcttgcgaggattaatagaagag
cggtatttattgtttattctcgaacaatctggtattaatgatgaaatgactgcacatcat
ttatcaaatcaacagtttcaaacttttattaatcttctaaagacttttacctttacagta
gatggcacacttccattagacaaagcttttgtgacaggcggcggtatttctttaaaagaa
atagaaccaaaaactatgatgtctaaattagtgccaggattattttttatgtggcgaagtc
ctagatattcatggatacactggaggttataatataacaagcgcattagttacgggtcat
gttgctggaatgtttgctggtgaatttaaaatagatcaaaacaaataa

Sequence 2092
MQLDVFPELNVDQLSQKVRKILNAEPDKYIKNSLRGLIEERYLLFILEQSGINDEMTAHH
LSNQQFQTFINLLKTFTFTVDGTLPLDKAFVTGGGISLKEIEPKTMMSKLVPGLFLCGEV
LDIHGYTGGYNITSALVTGHVAGMFAGEFKIDQNK*

Sequence 2093
Contig_0707_pos_1717_1337,
putative peptide of unknown function
atgataaatgcaaatttaagtgctagtactagagtgaaacaaaatgctcgtacatcgata
aatgaaattgttagtaacgctttaagtcaacttaataaagtaaccacaaataaagaagtt
gatgaaatagttaacgaaacgattgaaaaacttaagtcaatacaaataagagaagataaa
atattgagtagtcaacgttcatcaacatctatgacggaaaaatctaatcaatgttatagt
tccgagaataatacaattaaatctctaccagaggcaggaaatgctgataaatcactacca
ttagcaggagttactttaatatctggtttagcaatcatgtcctcacgtaaaaagaaaaaa
gataaaaaagtaaatgactaa

Sequence 2094
MINANLSASTRVKQNARTLINEIVSNALSQLNKVTTNKEVDEIVNETIEKLKSIQIREDK
ILSSQRSSTSMTEKSNQCYSSENNTIKSLPEAGNADKSLPLAGVTLISGLAIMSSRKKKK
DKKVND*

Sequence 2095
Contig_0707_pos_0_497,
putative peptide of unknown function
atggatattttttctttttttgattttctgctctttataaacaaactgactactggagat
gattcaatgaagacctataagccgtaccgacatcaattaaggcgttcgctatttgcctca
acgattttcccagtatttatggtgatgattattggtttaataagctttttatgctatttat
ataaggtcgaacattgcaccattcatcagcatacctatcaaactcaaaccgaattacaa
cgtatcgacaaacattttcatacgtttgttacgcagcaacaaaaacaatggcgtcatgtt

```
gatttatcacatccaactgatatcacaaagatgaaacgccaactattaaaacaagtccat
caacaacctgcgatattgtattacgatttaaaaggttcttcacaatctttcacaaacaat
tatgaacaattagacacaacaaagatgtatttaatatcaaaatatcgaattgattttaaa
gacgatacttatatcct
```

Sequence 2096
MDIFLFLIFLLFINKLTTGDDSMKTYKPYRHQLRRSLFASTIFPVFMVMIIGLISFYAIY
IWVEHCTIHQHTYQTQTELQRIDKHFHTFVTQQQKQWRHVDLSHPTDITKMKRQLLKQVH
QQPAILYYDLKGSSQSFTNNYEQLDTTKMYLISKYRIDFKDDTYIL

Sequence 2097
Contig_0708_pos_1581_3635,
is similar to (with p-value 0.0e+00)
>gp:gp|AF090142|AF090142_1 Staphylococcus epidermidis lipase
  precursor (gehD) gene, complete cds. NID: g3789931.

```
gtgatttctttgacaaataataatagaataaaaagatttagcattagaaaatacgcagtg
ggagtagtatcaattattacaggtgtaacaattttcatcggagggcaacaagctcaagca
gctgaaacttcagtgcaacatgcggatgcacacccagaagactcgcaaacaacacaacaa
ttaaaaaatgataaggtagaagaaacgttaaaagcttcaaaacaaggtactgcttatagt
caacaagtacaaacaattaatcaatctaaaacaaatcaaaacaaccaacattctgtagct
gaaagtgaacaactgaaaagtgatgagacagctaatcagccaaaaaaagaagaaggtagt
tcagtaaaaacaagacgtccaaccatctaaaaatgtaaatcaacaagacgcagctactcaa
tcaaatgagagaaaaaatattgacataaaaggtgaaggtcaaacttcaaagacaagcaat
caacatattcagagttctaacagtcataatcaatcaacagaaacaaaagacagcgactca
gaagaaatcgatcaaccattagtgaaattacaaaagccgtctaatgattctacatatcaa
acacaatcaaaagctaaacaagatagttctaaacagctccctcaagaaaaaacaacaaaa
cgtcaaatccaaacaactgaaaatgaacagacaactaaagttgattctaaaaaagctaat
gacactcaaaatgttgaacaacatactcaagagcctaaaaatgatacatcaacatcacaa
aaaaatcatcatcaagtagctacaaaagaacaatctaatagaagtacaacaagggagacg
caaagcaatcagcaaatgctaatcaaaatcatcagtctacacatcaagcacagttcaaa
aaccaatatccagtagtatttgtccatggtttcctaggctttgcaggtgataatcaattt
agtttagctccaaaatattggggtggtacaaaatacaatattgacagaaatttaactaat
gagggatacaatgtacatgaggcaaatattggtgctttagtagtaactatgatcgcgca
gtagaattgtattactatgtcaaaggaggacgtgttgattacggtgcagcgcatgcagct
aaatatggtcatcatcgctatggtcgaacatacaagggcatcatgcgtgattgggaacct
ggcaaaaaaattcattttataggtcacagtatgggtggtcaaaccattcgtcaaatggaa
gaattcttaagaaatggtaaccaagaagaaatagaatatcaacgtcaacatgggggtact
atatccgatttatttacaggtggtaaagataatatggttgcttcaattactacacttggc
acaccacataatggtacacctgctgcagataaaattggcacacgtaaacttgtaaaagaa
acgattaatcgtattggtagattaagtggtggtaaagatgtagatatagatttaggtttt
tctcaatggggattaaaacaacaaccaaatgaaagctacattgattacgcggaacgtgta
tccaaaagtaagatttggaatactgaagatcaagctgttaatgatctgacaacgcaaggt
gctgaaaaaattaatcaacaaacaagtctaaatcctaatattgtctacactacttataca
gggtcagcgactcacacaggacctcttggtaatgaattacctaattctagtgaaattctt
ttgttgaacttaaccagccgtattattggtaaagatgcaaacaaagaaattagaccgaat
gatggtgtagttccagttatatcatcacaacatccttctaatcaagcctttaaaaaagtt
gatgatcatacaccagctactgataaaggtgtttggcaagtgagacccgttcaacatgat
tgggaccatttagatttagtaggtatggatgcatttgatttaacacatacaggtagagaa
ttaggtcaattctatctaggtattatggataatatcatgagaatcgaagaagcagaaggt
attacaatataaataa
```

Sequence 2098
VISLTNNNRIKRFSIRKYAVGVVSIITGVTIFIGGQQAQAAETSVQHADAHPEDSQTTQQ
LKNDKVEETLKASKQGTAYSQQVQTINQSKTNQNNQHSVAESEQLKSDETANQPKKEEGS
SVKQDVQPSKNVNQQDAATQSNERKNIDIKGEGQTSKTSNQHIQSSNSHNQSTETKDSDS
EEIDQPLVKLQKPSNDSTYQTQSKAKQDSSKQLPQEKTTKRQIQTTENEQTTKVDSKKAN
DTQNVEQHTQEPKNDTSTSQKNHHQVATKEQSNRSTTRETQKQSANANQNHQSTHQAQFK
NQYPVVFVHGFLGFAGDNQFSLAPKYWGGTKYNIDRNLTNEGYNVHEANIGAFSSNYDRA
VELYYYVKGGRVDYGAAHAAKYGHHRYGRTYKGIMRDWEPGKKIHFIGHSMGGQTIRQME

522

EFLRNGNQEEIEYQRQHGGTISDLFTGGKDNMVASITTLGTPHNGTPAADKIGTRKLVKE
TINRFGRLSGGKDVDIDLGFSQWGLKQQPNESYIDYAERVSKSKIWNTEDQAVNDLTTQG
AEKINQQTSLNPNIVYTTYTGSATHTGPLGNELPNSSEILLLNLTSRIIGKDANKEIRPN
DGVVPVISSQHPSNQAFKKVDDHTPATDKGVWQVRPVQHDWDHLDLVGMDAFDLTHTGRE
LGQFYLGIMDNIMRIEEADGITNK*

Sequence 2099
Contig_0710_pos_168_1115,
is similar to (with p-value 7.0e-45)
>gp:gp|Z99108|BSUB0005_69 Bacillus subtilis complete genome
(section 5 of 21): from 802821 to 1011250. NID: g2633055. >g
p:gp|D78509|D78509_11 Bacillus subtilis YfjG-YfjR genes, com
plete cds. NID: g2780390.
atgacggtgaccgttagatttcaatctttatcgcaacctcttacattagtttcaaatgtg
aaagagattcctaaagatgcaacgattatatggtatgattttgaaaatgccactgatgaa
gaaaatgagtatttaaaaaatcattttgatttcaattacttagaaatagatgatgctatc
aatggtgacccacgagttaaatatatagaatatgacgcgtatcaatatatgatatttcat
agtattattaatgatgattactcaccaatctcactaagtgtatttttagaaggtaatgtt
ttagtgacataccatcacaaacattttccatcattaaagcgtgtggctcaatacaatgca
gaaaatcatgatagtgaattagattgtgcagacatcgtcattcatattctggattgtatg
gtggataaatattttaactttgtttatggtattgaagataaagtgtataattttgaagct
aagcatgtcgatgaccgctatagtaagagcgttatggaaaatgtctttcaattacgttcg
gatttaattaaaaatcaaacgcgtattatttccgatgcaagaagttgtagatacaatgaaa
caagaaggaaatataattaaagatgccaaacatagaatgtatattcaacatattgatgat
catcttattaaacaaagaagtgttattcggacttctcaagaaatgacgaatgagattcgt
gaaaattatgaatcattcacctcatttaggatgaatagtataatgcagatacttacgctt
gtatctgttatattctcaccactcactttattgctggtgtatatggaatgaactttgaa
tttatgcctgagttgaaatggcattatgcttatttcgtgtgcttaactttaatgctaatt
ataacaataatattaatcatattctttaaaaagaaaaaatggttttaa

Sequence 2100
MTVTVRFQSLSQPLTLVSNVKEIPKDATIIWYDFENATDEENEYLKNHFDFNYLEIDDAI
NGDPRVKYIEYDAYQYMIFHSIINDDYSPISLSVFLEGNVLVTYHHKHFPSLKRVAQYNA
ENHDSELDCADIVIHILDCMVDKYFNFVYGIEDKVYNFEAKHVDDRYSKSVMENVFQLRS
DLIKIKRVLFPMQEVVDTMKQEGNIIKDAKHRMYIQHIDDHLIKQRSVIRTSQEMTNEIR
ENYESFTSFRMNSIMQILTLVSVIFSPLTFIAGVYGMNFEFMPELKWHYAYFVCLTLMLI
ITIILIIFFKKKKWF*

Sequence 2101
Contig_0710_pos_1206_1763,
is similar to (with p-value 2.0e-29)
>gp:gp|AF016485|AF016485_60 Halobacterium sp. NRC-1 plasmid
pNRC100, complete plasmid sequence. NID: g2822278. >gp:gp|AF
016485|AF016485_149 Halobacterium sp. NRC-1 plasmid pNRC100,
complete plasmid sequence. NID: g2822278.
atgtcacaaaaagatgcgctggtttcagattttgataaagtgagatttgttcatcattcc
atccccagtattgatgttagtcaagtcgatatgacaagtcatactacgaaattcgatttg
gcatatccaatctatataaatgcaatgactggtggaagtgattggacaaaacaaattaat
gaaaaattagcaattgttgctagagaaactggaattgcaatggcggtgggatcaacacat
gcagctttgcgcaatcctaatatgattgaaacatttagcattgtgcgtaaaacaaatccc
aaaggaacaattttcagcaatgtgggtgccgatgtaccagtggataaagctctacaagcg
gttgaattattagatgctcaagcgctacaaattcatgtgaactcacctcaagaattagtc
atgcctgaagggaaccgtgaatttgcttcatggatgtcaaatattgaatctattgttaaa
cgcgttgatgttccagttattattaaagaagttggtttcggaatgagtaaagaaacatta
caagcgttatatgattaa

Sequence 2102
MSQKDALVSDFDKVRFVHHSIPSIDVSQVDMTSHTTKFDLAYPIYINAMTGGSDWTKQIN
EKLAIVARETGIAMAVGSTHAALRNPNMIETFSIVRKTNPKGTIFSNVGADVPVDKALQA

VELLDAQALQIHVNSPQELVMPEGNREFASWMSNIESIVKRVDVPVIIKEVGFGMSKETL
QALYD*

Sequence 2103
Contig_0710_pos_1830_3305,
is similar to (with p-value 4.0e-90)
>gp:gp|AF006665|AF006665_31 Bacillus subtilis 168 region at
182 min containing the cge gene cluster. NID: g2529445. >gp:
gp|AF015775|AF015775_7 Bacillus subtilis YodA (yodA), YodB (
yodB), YodC (yodC), YodD (yodD), ABC-transporter (yodE), per
mease (yodF), proteinase (ctpA), YodH (yodH), YodI (yodI), c
arboxypeptidase (yodJ), purine nucleoside phosphorylase (deo
D), YodL (yodL), YodM (yodM), YodN (yodN), YodO (yodO), YodP
 (yodP), acetylornitine deacetylase (argE), butirate-acetoac
etate CoA transferase (yodR), butyrate acetoacetate-CoA tran
sferase (yodS), YodT (yodT), CgeE (cgeE), CgeD (cgeD), CgeC
(cgeC), CgeA (cgeA), CgeB (cgeB), YzxA (yzxA), UDP-glucose e
pimerase (yodU), YodV (yodV), and YodW (yodW) genes, complet
e cds; and YodZ (yodZ) gene, partial cds. NID: g2415383. >gp
:gp|Z99114|BSUB0011_121 Bacillus subtilis complete genome (s
ection 11 of 21): from 2000171 to 2207900. NID: g2634230.
atgaatgatcatcaaaaaaatcatgcaacatctcaagatgataacacaatgtcaacacca
tctaagaatagcaagcatataaaaattaaattatggcatttcatactcgttattttgggt
attattcttttaacatccatcattactgtagtatcaacaattttaattagccatcaaaaa
agtggtttaaataaagaacaacgtgcaaatttaaaaaaaattgaatatgtctatcaaaca
cttaataaagattattacaaaaagcaaagttctgataaattaactcaatctgccatagat
ggtatggttaaagaacttaaagatccatattcagaatatatgactgctgaagaaacaaaa
caatttaatgaaggtgtatcaggtgatttcgttggcataggtgctgaaatgcaaaagaaa
aatgaacagataagtgttactagcccaatgaaggattcaccagcagaaaaagctggtatt
caacctaaagatatcgtcacacaagtgaatcatcattcggtagtcggtaaaccacttgat
caagttgttaaaatggtccgcggcaaaaaaggaacatatgttactttaactataaaacgt
ggttcgcaagaaaaggatattaagattaaacgcgataccattcacgttaagagtgtagag
tatgagaagaaaggcaatgtaggcgtactaacaatcaataaattccaaagcaatacttct
ggtgaactcaaatctgcaatcatcaaagcgcataagcaaggcatccgtcatatcatttta
gatttgagaaataatccggggggggttattagatgaggcagtcaagatggctaacatcttt
attgataagggaaatactgtcgttcaattagaaaaaggtaaggataaggaagaattaaaa
acttctaatcaagcactaaaacaagcaaaagatatgaaagtatccatcttagttaatgag
ggatcagctagtgcttcagaagtgtttacaggtgctatgaaagactatcataaagctaaa
gtttacggttctaaaacatttggtaaaggtatcgttcagaccactcgtgaatttagtgat
ggttcattaattaaatatacagagatgaaatggctaacgcctgatggccattatattcat
ggtaaaggaattagaccagatgttagtatctcaacaccaaaataccaatcactcaatgtc
attccagataacaaaacttatcatcaaggtgaaaaagataaaaatgttaaaacgatgaaa
ataggtctaaaagctttaggttatccaattgataacgaaacaaacatatttgacgaacaa
ttagaatctgctattaaaacatttcaacaagacaataatttaaaagttaatggcaatttt
gataaaaaaacaaatgataaatttactgaaaaactagttgaaaaagcgaataaaaaagat
actgttttaaacgatttactaaacaaactaaaataa

Sequence 2104
MNDHQKNHATSQDDNTMSTPSKNSKHIKIKLWHFILVILGIILLTSIITVVSTILISHQK
SGLNKEQRANLKKIEYVYQTLNKDYYKKQSSDKLTQSAIDGMVKELKDPYSEYMTAEETK
QFNEGVSGDFVGIGAEMQKKNEQISVTSPMKDSPAEKAGIQPKDIVTQVNHHSVVGKPLD
QVVKMVRGKKGTYVTLTIKRGSQEKDIKIKRDTIHVKSVEYEKKGNVGVLTINKFGNTS
GELKSAYLKAHKQGIRHIILDLRNNPGGLLDEAVKMANIFIDKGNTVVQLEKGKDKEELK
TSNQALKQAKDMKVSILVNEGSASASEVFTGAMKDYHKAKVYGSKTFGKGIVQTTREFSD
GSLIKYTEMKWLTPDGHYIHGKGIRPDVSISTPKYQSLNVIPDNKTYHQGEKDKNVKTMK
IGLKALGYPIDNETNIFDEQLESAIKTFQQDNNLKVNGNFDKKTNDKFTEKLVEKANKKD
TVLNDLLNKLK*

Sequence 2105

Contig_0710_pos_4812_3682,
is similar to (with p-value 7.0e-51)
>sp:sp|P53434|YRP2_LISMO HYPOTHETICAL 41.4 KD PROTEIN IN RPO
D 2'REGION (ORFA2). >gp:gp|U17284|LMU17284_3 Listeria monocy
togenes major sigma factor (rpoD) gene, partial cds, and dow
nstream orfA1 and orfA2 genes, complete cds. NID: g687597.
atggttggtaatgtaggattattaattggtaatgataagttagatattacaggtattctg
acaacactcgactgcaccgatgatgttgttaaccaagcaattgaacttaataccaatacc
atcattgctcatcatccacttattttcaaaggagtaaaacgtatcgttgaagatggatat
ggtagtataattcgtaaacttatccaaaataatatcaatcttatagcattacacactaat
cttgatgtaaatcctaaaggtgtcaatcgaatgttagcggatcaaataggtttagagaac
atatcaatgattaatacaaatagctcatattattacaaagttcaaacttttatacctaaa
aattatattgaagatttcaaagacagtttaaacgaacttggattagctaaagaaggtaat
tacgaatattgtttctttgaaagtgaaggtaaagggcaatttaaaccagtaggtgatgca
agtccttatataggggaagttagatagtatcgaatatgttgatgaaataaaaacttgagttt
atgataaaagacaatgaattagaaataactaaacgtgctattttagataatcacccatac
gaaacaccagtttttgattttattaaaatgaacaaagaaagtgagtatggattagggatt
attggacaattaaaccaactatgactttagatgaattttctgaatatgccaaaaaacag
ctcaatataccgagcgtacgatatacaggtcaacatgatagtccaattaagaaagtagct
atcataggtggttcaggtataggatttgagtataaagctagccaacttggagcagatgtt
tttgttactggtgatattaaacaccatgatgctttagatgctaaaatccaaaatgtaaat
ttattagacatcaatcattatagtgagtatgttatgaaagaaggattaaaagaattatta
gaaaaatggttatttaaatatgaaaatcaatttccaatatatgcttctgaaatcaacaca
agtgtttatttttcgatgtcttcttcaaaatgctctttcgcgggatattcatcatcccat
tggttgttgttgaattcttccataatctgtttagcctctttgactaattga

Sequence 2106
MVGNVGLLIGNDKLDITGILTTLDCTDDVVNQAIELNTNTIIAHHPLIFKGVKRIVEDGY
GSIIRKLIQNNINLIALHTNLDVNPKGVNRMLADQIGLENISMINTNSSYYYKVQTFIPK
NYIEDFKDSLNELGLAKEGNYEYCFFESEGKGQFKPVGDASPYIGKLDSIEYVDEIKLEF
MIKDNELEITKRAILDNHPYETPVFDFIKMNKESEYGLGIIGQLNQTMTLDEFSEYAKKQ
LNIPSVRYTGQHDSPIKKVAIIGGSGIGFEYKASQLGADVFVTGDIKHHDALDAKIQNVN
LLDINHYSEYVMKEGLKELLEKWLFKYENQFPIYASEINTSVLFSMSSSKCSFAGYSSSH
WLLLNSSIICLASLTN*

Sequence 2107
Contig_0711_pos_2494_2901,
putative peptide of unknown function
atgaaattatctattttttcatgatggtcaatttttttgtaggtgtcgttgaataccaagaa
ggcttcattcataaatatctaaaagttacatttggcaatgaacctagcgatgaaacagtg
ttacgattcataacttttaaacttattcctttattaaatcaaacacacggtaagaagaaa
cctattcaaaagcataaaaagattaatccaaaacgtttacaacgtaaatcgctaaagaa
caaaaagagaccaatttaactacatttgctcaacaagcgattaaagaagaacaagagttg
aataagctaaagagtaaaaaacttcagcgattagaaaaagaacgacacagacaatacaaa
agaatgttaaaaagaaaaaaagcacatgaaaagcacaaaggtcactaa

Sequence 2108
MKLSIFHDGQFFVGVVEYQEGFIHKYLKVTFGNEPSDETVLRFITFKLIPLLNQTHGKKK
PIQKHKKINPKRLQRKIAKEQKETNLTTFAQQAIKEEQELNKLKSKKLQRLEKERHRQYK
RMLKRKKAHEKHKGH*

Sequence 2109
Contig_0711_pos_3571_3053,
putative peptide of unknown function
atgttagaaacacatagattaaagctagtgaagcctaatttgagttatacagatgaactt
tatcaattgcatacaaataaggtagctacaaagtatacacctaaaggtattcatcagaat
aaagtagcaacccaagatttattaaaggatggatgaggcattgggatgaatatcaattt
ggttacttcattttaattatgagagataatcacgaagtagtggggatagcgggatttgag
tatcgtacaattcatcaacaacagtttcttaatgcgtattatagaatctttccatcgtat

actggtgttggtttagcttttgagtcaatggaggagattgcccgtcatttaaaaaagcat
gataccataacaccaaaattaattcgaacaaatcaatataatacaaattctattaaatta
gcacaaaaactcggatataattatgatgctaactgggacgatgtaattaataaaggagat
cgttgttttttttaacctacaagcgttggataataactaa

Sequence 2110
MLETHRLKLVKPNLSYTDELYQLHTNKVATKYTPKGIHQNKVATQDFIKGWMRHWDEYQF
GYFILIMRDNHEVVGIAGFEYRTIHQQQFLNAYYRIFPSYTGVGLAFESMEEIARHLKKH
DTITPKLIRTNQYNTNSIKLAQKLGYNYDANWDDVINKGDRCFFNLQALDNN*

Sequence 2111
Contig_0711_pos_703_359,
putative peptide of unknown function
atggagatgatagaagaacgtaatttatcagggcttattcaaacactaactttcaatcat
cccatcattcaaattcttaaagagaacacattaaatcaacttaaaatactctctcattat
ttaccagagcgacaccctgcaatggtggcaattcaatcttggtcacaatggtttatgat
catgggattactgaaatccaccttgatgtaactgcacaagcgcctagatcttattacaaa
ggtattttttataaaatgtcatcttaaaaatactgctcatagcgttttgacaggtggatat
tatcacggttcactagaaggtttgtttaggattaacactttaa

Sequence 2112
MEMIEERNLSGLIQTLTFNHPIIQILKENTLNQLKILSHYLPERHPAMVAIQSWSQWFTD
HGITEIHLDVTAQAPRSYYKGIFIKCHLKNTAHSVLTGGYYHGSLEGFGLGLTL*

Sequence 2113
Contig_0711_pos_0_334,
putative peptide of unknown function
atgttacgagttgcattagcaaagggtcgtttattaaagagtttttatcgaatatttacaa
caagttaatcagatagatattgcaactgtacttttaaatagacagcgacagttattgctt
acagtcgacaacattgaaatgattttagttaaaggaagcgatgtgcctacttatgtagaa
caaggtattgctgatgtaggaatagtgggaagtgatattctgaatggtcaaaaatataat
attaataaattactcgatttgccatttggtaaatgtcattttgcgttggcggcaaagcca
gaaacatctcgctataaaaaagtagcaaGTATTG

Sequence 2114
MLRVALAKGRLLKSFIEYLQQVNQIDIATVLLNRQRQLLLTVDNIEMILVKGSDVPTYVE
QGIADVGIVGSDILNGQKYNINKLLDLPFGKCHFALAAKPETSRYKKVASIX

Sequence 2115
Contig_0712_pos_299_928,
is similar to (with p-value 1.0e-52)
>sp:sp|Q02499|KPYK_BACST PYRUVATE KINASE (EC 2.7.1.40) (PK).
 >pir:pir|S29783|S29783 pyruvate kinase (EC 2.7.1.40) (versi
on 2) - Bacillus stearothermophilus >gp:gp|D13095|BACPK_3 B.
 stearothermophilus phosphofructokinase and pyruvate kinase
genes. NID: g285620.
atgccttttctaaacctataattgcgcttggtgaagtaataccattttcttctgtaatt
agacctatagcttttttcaacatatggtactaatgtttcatcaacagaatttgtaataata
actttatcagataaatctttaccttctaaatcactagcactatctgcgacaattgcatgg
cctacaacagatcctctaccaacaccttggcctttagcaatctcatcacctactaagtgg
attttcatcatatttgtagttccttttttctccagtaggtacaccagcagtaataataatt
aaatctccgtttgaaactctaccagtttctactgctgttgctacagcattatttagtaaa
gcatcagttgttttacgtccttctttaacgaccggatttactccccatacaattgcacat
tgtctagcagttttttcgctaggtgttacagcaatgatatctgaatgtggacgatattta
gaaatcgtacgtgctgttgaaccactttcagttgctgctacaatagctttttacattaag
tttaaggcagtatgtgcaacagaaacaccaattgcatttactaatgaagtttcaactaat
ttagtacgatcacttaataatttttttatag

Sequence 2116

MPFSKPIIALGEVIPFSSVIRPIAFSTYGTNVSSTEFVIITLSDKSLPSKSLALSATIAW
PTTDPLPTPWPLAISSPTKWIFIIFVVPFSPVGTPAVIIKSPFETLPVSTAVATALFSK
ASVVLRPSLTTGFTPHTIAHCLAVFSLGVTAMISECGRYLEIVRAVEPLSVAATIAFTFK
FKAVCATETPIAFTNEVSTNLVRSLNNFL*

Sequence 2117
Contig_0712_pos_5629_0,
is similar to (with p-value 2.0e-21)
>gp:gp|Z99125|MLCL536_19 Mycobacterium leprae cosmid L536. N
ID: g2398683. >gp:gp|U00013|U00013_5 Mycobacterium leprae co
smid B1496. NID: g466868. >gp:gp|Z99125|MLCL536_19 Mycobacte
rium leprae cosmid L536. NID: g2398683.
atgaattttaataatttggatcaattatatagatctgtaattatggatcattacaaaaac
cctagaaacaaaggtgtcttagacaatggctcaatgactgttgatatgaataaccctaca
tgtggtgatcgcatacgtttgacatttgatattgaagacggaatcattaatgatgctaag
tttgaaggagaaggatgttcaatttcaatgtctagtgcatctatgatgactgaagcagtt
aaaggtcattcacttggtgaagcaatgcaaatgagccaagagtttactaaaatgatgctc
ggtgaagactacgagattacagaagaaatgggagatattgaggcgcttcaaggtgtctca
caattcccagctagaattaaatgtgcaacgcttgcatggaaagcattagaaaaagggaca
gtcgaaaaagaaggtaagtcagaaggtaca

Sequence 2118
MNFNNLDQLYRSVIMDHYKNPRNKGVLDNGSMTVDMNNPTCGDRIRLTFDIEDGIINDAK
FEGEGCSISMSSASMMTEAVKGHSLGEAMQMSQEFTKMMLGEDYEITEEMGDIEALQGVS
QFPARIKCATLAWKALEKGTVEKEGKSEGT

Sequence 2119
Contig_0712_pos_5603_5016,
is similar to (with p-value 3.0e-27)
>sp:sp|P16468|MAOX_BACST MALATE OXIDOREDUCTASE (NAD) (EC 1.1
.1.38) (MALIC ENZYME). >pir:pir|A33307|DEBSXS malate dehydro
genase (oxaloacetate-decarboxylating) (EC 1.1.1.38) - Bacill
us stearothermophilus >gp:gp|M19485|BACMAL_1 B.stearothermop
hilus malic acid gene, complete cds. NID: g143164.
atgcttgaactaattgatctacatcttctttcgtattataaacgtaaaaacttgcacgtg
ctgttgaagagacacctaaccatttcatcaaaggttgcgcacagtgatggcctgcacgaa
ctgcaacaccttctaggtgtatctgtagcggatatattaactcaagatatggtgaagaca
atggctgatgatccaattattttttgctatggctaatcctaatcctgaaatcaatcccaat
gaagccaaacaggcaggtgcaaaggtcgtaggtacaggtcgatctgattttccaaaccaa
attaataatgtattagcatttccaggtattttagaggtgcattggatgttgaagccact
catattaacgaagatatgaaaaaggcagctgtagaagctatcgttcatttaatagacgaa
aatgagttacatcctgattactgtataccaggaccatttgataaaagagtagctccatca
gtagccaaaaatgtagctaaagctgctatggaatccggtgtagcaagaattaaaattgat
acacaagaaatatttgataaaactatgaaacttactgacttaaaatag

Sequence 2120
MLELIDLHLLSYYKRKNLHVLLKRHLTISSKVAHSDGLHELQHLLGVSVADILTQDMVKT
MADDPIIFAMANPNPEINPNEAKQAGAKVVGTGRSDFPNQINNVLAFPGIFRGALDVEAT
HINEDMKKAAVEAIVHLIDENELHPDYCIPGPFDKRVAPSVAKNVAKAAMESGVARIKID
TQEIFDKTMKLTDLK*

Sequence 2121
Contig_0712_pos_2893_1970,
is similar to (with p-value 0.0e+00)
>sp:sp|P00512|K6PF_BACST 6-PHOSPHOFRUCTOKINASE (EC 2.7.1.11)
(PHOSPHOFRUCTOKINASE) (PHOSPHOHEXOKINASE). >pir:pir|A27474|
KIBSFF 6-phosphofructokinase (EC 2.7.1.11) - Bacillus stearo
thermophilus >gp:gp|M15643|BACPFK_1 B.stearothermophilus 6-p
hosphofructo-1-kinase gene, complete cds. NID: g143311. >gp:

gp|D13095|BACPK_2 B. stearothermophilus phosphofructokinase
and pyruvate kinase genes. NID: g285620.
atgaacgctgctgttcgtgcagtgactcggacagcaatttacaataatattgaagtttat
ggtgtttatcaaggttaccaaggtttacttgatgatgatattcataagcttgaattgggt
tcagtaggggatacaattcaacgaggaggaactttcctcttttccgcaaggtgccctcag
ttcaaagaagaggatgtacgtaagaaagctattgagaatttacgtaagcgtggtatcgaa
ggtttagttgttattggaggagatggcagctatagaggggcacaacgaattagtgaggaa
tgtaaagaaattcaaacaattggtattcctggtacaattgataatgatattaatggtaca
gattttacaattggttttgatactgcattaaacactattattgaatcagtcgataagatt
agagatacggcatcaagtcacgcaagaacgtttattgttgaagttatggggcgtgattgt
ggagatttagctttatgggctggattatctgtaggtgctgaaacgattgtttttaccagaa
gtcaatacagatattaaggatgtagctgaaaagattgaacagggtattaaaagagggaaa
aaacattctatcgttatggttgcagaaggttgtatgagcggccaagaatgtgcagatgag
ttaacgaagtatattaacattgatacacgagtttcagtgttaggtcacattcaacgtggc
ggtagcccatctggtgctgatcgagtattagcttctcgacttggtggatatgctgttgaa
ctattaaaacaaggcgagacagctaaaggtgttggcattaggaataatcaattaacctct
acgccgtttgatgaaatttttgctgaaagtgatcgcaaatttaatagtcaaatgtatgaa
ttagcaaaagaattatcaatttaa

Sequence 2122
MNAAVRAVTRTAIYNNIEVYGVYQGYQGLLDDDIHKLELGSVGDTIQRGGTFLFSARCPQ
FKEEDVRKKAIENLRKRGIEGLVVIGGDGSYRGAQRISEECKEIQTIGIPGTIDNDINGT
DFTIGFDTALNTIIESVDKIRDTASSHARTFIVEVMGRDCGDLALWAGLSVGAETIVLPE
VNTDIKDVAEKIEQGIKRGKKHSIVMVAEGCMSGQECADELTKYINIDTRVSVLGHIQRG
GSPSGADRVLASRLGGYAVELLKQGETAKGVGIRNNQLTSTPFDEIFAESDRKFNSQMYE
LAKELSI*

Sequence 2123
Contig_0712_pos_1946_189,
is similar to (with p-value 0.0e+00)
>sp:sp|P51181|KPYK_BACLI PYRUVATE KINASE (EC 2.7.1.40) (PK).
 >pir:pir|JC4220|JC4220 pyruvate kinase (EC 2.7.1.40) - Baci
llus licheniformis >gp:gp|D31955|BACPYK2_2 Bacillus lichenif
ormis gene for pyruvate kinase, complete cds. NID: gl041098.
atgagaaagactaaaattgtatgtacaataggaccagcttcagaatcagaggaaatgctc
gaaaaactaatgaatgcaggaatgaacgttgcgcgttttaaatttctcacatggtagtcat
gaagaacataaagcaagaattgatacaattcgtaaagttgctaaacgtttaaataaaaca
attggcttgttattggatactaaaggggccagaaattcgtacgcacaatatgaaagatgga
cttattgtttttagaaaaaggcaaagaagtcattgtcagtatgaatgaagttgaaggaaca
cctgaaaaattctctgtaacatatgaaaatctaatcaatgatgtcaatattggatcatat
atactattagatgatggtttagttgaacttcaagtcaaagaaattaacaaagataaaggc
gaagttaaatgtgatatcttaaatactggtgaattaaaaaataaaaaaggtgttaactta
cctggtgttaaagttaatttacctggtatcactgataaagatgccgatgatatcagattt
ggtataaaggaaatgtagactttatagctgcaagttttgtaagacgtccaagtgatgtt
ttagatatccgtcaaattcttgaagaagaaaaagcagaaataacaattttccctaaaatc
gaaaaccaagaaggtatcgataatattgaagaaattcttgaagtatctgatggattaatg
gtagcacgtggtgatatgggtgttgaaattccaccagaaagcgtaccaatggttcaaaaa
gatttaattagaaaatgtaataaattaggaaaacctgtaattactgcgactcaaatgctt
gattctatgcaacgtaatccacgtgcgacacgtgcagaagcaagtgacgtagctaatgca
atatacgatggtactgacgctgtaatgttatcaggcgaaactgcagcaggtcaatatcct
gaagaagctgttaaaactatgcgtaatattgcagtttctgctgaagcagcgcaagactat
aaaaaattattaagtgatcgtactaaattagttgaaacttcattagtaaatgcaattggt
gtttctgttgcacatactgccttaaacttaaatgtaaaagctattgtagcagcaactgaa
agtggttcaacagcacgtacgatttctaaatatcgtccacattcagatatcattgctgta
acacctagcgaaaaaactgctagacaatgtgcaattgtatggggagtaaatccggtcgtt
aaagaaggacgtaaaacaactgatgctttactaaataatgctgtagcaacagcagtagaa
actggtagagtttcaaacggagatttaattattattactgctggtgtacctactggagaa
aaaggaactacaaatatgatgaaaatccacttagtaggtgatgagattgctaaagcccaa
ggtgttggtagaggatctgttgtaggccatgcaattgtcgcagatagtgctagtgattta

gaaggtaaagatttatctgataaagttattattacaaattctgttgatgaaacattagta
ccatatgttgaaaaagctataggtctaattacagaagaaatggtattacttcaccaagc
gcaattataggtttagaaaaaggcatacctactgttgttggtgtagaacaagcaactaaa
gaaattaaaaatgatatgttagtgactttagatgcgtcacaaggtaaagtgtttgaaggt
5   tatgctaacgtcctttaa

Sequence 2124
MRKTKIVCTIGPASESEEMLEKLMNAGMNVARLNFSHGSHEEHKARIDTIRKVAKRLNKT
IGLLLDTKGPEIRTHNMKDGLIVLEKGKEVIVSMNEVEGTPEKFSVTYENLINDVNIGSY
10   ILLDUGLVELQVKEINKDKGEVKCDILNTGELKNKKGVNLPGVKVNLPGITDKDAIDIRF
GIKENVDFIAASFVRRPSDVLDIRQILEEEKAEITIFPKIENQEGIDNIEEILEVSDGLM
VARGDMGVEIPPESVPMVQKDLIRKCNKLGKPVITATQMLDSMQRNPRATRAEASDVANA
IYDGTDAVMLSGETAAGQYPEEAVKTMRNIAVSAEAAQDYKKLLSDRTKLVETSLVNAIG
VSVAHTALNLNVKAIVAATESGSTARTISKYRPHSDIIAVTPSEKTARQCAIVWGVNPVV
15   KEGRKTTDALLNNAVATAVETGRVSNGDLIIITAGVPTGEKGTTNMMKIHLVGDEIAKGQ
GVGRGSVVGHAIVADSASDLEGKDLSDKVIITNSVDETLVPYVEKAIGLITEENGITSPS
AIIGLEKGIPTVVGVEQATKEIKNDMLVTLDASQGKVFEGYANVL*

Sequence 2125
20   Contig_0714_pos_181_783,
is similar to (with p-value 2.0e-33)
>gp:gp|D50453|D50453_106 Bacillus subtilis DNA for 25-36 deg
ree region containing the amyE-srfA region, complete cds. NI
D: g1805369.
25   atgcttccaaagctgtttacgttggtaaagatttatataaaaggtaagtcataccttcct
ttcgtaaatcgatctgctgtagcaggtattttaacaactggtgtcatgcgcaccttattg
tttttagctgtactaggtgttgttgtaactggcgttacgcttagttcagaaaatccacca
gcatcagttttccaacatgcattaggtcctataggtaaaaatattttttggcgtagtaata
tttgcagcagcaatgtcctcagtaattggttctgcatatacaagcgcaacatttttaaaa
30   acactacacaaatcgttactcaataaaaataatcttatcgttattacatttattgaaatt
tcaacttttgtttttcttatttattggtaaaccggtgagtttacttataatagctggtgcg
attaatggttggattctaccaatcacattaggtgcaattctcattgcaagtaggaaaaaa
tctatcgttggtaattaccaacacccaacatggatgcttgttttttggtattatagccgta
attgtcacaataatgactggtatctttttcattacaagatttagcaagtctttggaaaggt
35   taa

Sequence 2126
MLPKLFTLVKIYIKGKSYLPFVNRSAVAGILTTGVMRTLLFLAVLGVVVTGVTLSSENPP
ASVFQHALGPIGKNIFGVVIFAAAMSSVIGSAYTSATFLKTLHKSLLNKNNLIVITFIVI
40   STFVFLFIGKPVSLLIIAGAINGWILPITLGAILIASRKKSIVGNYQHPTWMLVFGIIAV
IVTIMTGIFSLQDLASLWKG*

Sequence 2127
Contig_0714_pos_1712_2215,
45   is similar to (with p-value 4.0e-46)
>sp:sp|P54452|YQEG_BACSU HYPOTHETICAL 20.1 KD PROTEIN IN NUC
B-AROD INTERGENIC REGION. >gp:gp|D84432|BACJH642_91 Bacillus
  subtilis DNA, 283 Kb region containing skin element. NID: g
262’063. >gp:gp|Z99117|BSUB0014_48 Bacillus subtilis complet
50   e genome (section 14 of 21): from 2599451 to 2812870. NID: g
2634966.
atgccaaatgcatatgtgaaatcaatatttgaaattgatatagaaaaacttgccgatagt
ggtgttaaaggtatcataactgatttagataatacacttgttggttgggatgttaaagaa
cctactaagggtgttaaatcatggtttgctaaggctaaagatttaggaataactgtcaca
55   attgtgtcaaataataataaaagtcgagtatcaagtttctcaagtaatttaggtgtagat
tatatattcaaagcacgtaaaccgatggggaaagcctttaagatggctattaaaaaaatg
aaaattcaaccgagagaaaccgttgttgtaggagatcaaatgcttactgatgtgtttggt
ggcaattgtaatggtttatatacaattatggtagtacctgttaaacggactgatggatta
attacaaagtttaatcgattaattgaaagacgattattaaatcattttagaaaaaaaggt

tatattaaatgggaggaaaattga

Sequence 2128
MPNAYVKSIFEIDIEKLADSGVKGIITDLDNTLVGWDVKEPTKGVKSWFAKAKDLGITVT
IVSNNNKSRVSSFSSNLGVDYIFKARKPMGKAFKMAIKKMKIQPRETVVVGDQMLTDVFG
GNCNGLYTIMVVPVKRTDGLITKFNRLIERRLLNHFRKKGYIKWEEN*

Sequence 2129
Contig_0714_pos_2822_3316,
is similar to (with p-value 6.0e-39)
>sp:sp|P54453|YQEH_BACSU HYPOTHETICAL 41.0 KD PROTEIN IN NUC
B-AROD INTERGENIC REGION. >gp:gp|D84432|BACJH642_92 Bacillus
 subtilis DNA, 283 Kb region containing skin element. NID: g
2627063. >gp:gp|Z99117|BSUB0014_47 Bacillus subtilis complet
e genome (section 14 of 21): from 2599451 to 2812870. NID: g
2634966.
atgatagatattccattagacgaaaaatcatttatgtttgatacaccaggtatcattcaa
tcacatcaaatgacaaattatgtatatgaaaatgagttgaaaatcattatacctaaaaat
gaaataaagcaacgtgtgtatcaacttaatgaaaaacagacattattttttcggaggattg
gcacgcattgattatgtatctggtggtaaaagaccacttgtttgtttcttttcaaatgat
ttaaatattcatagaactaaaaccgagaaagctaatgatttatggaaatcccaattaggc
gcattgctttcaccgcctcaagatgcacaacaatttaatcttaatgatgtaaaagcagta
agactggaaactggtaaaactaaacgtgacatcatgatatctggtttaggattcataact
attgatgctggtgcaaaagtgatagttcgtgttccaaaacatgtagatgttattttaaga
aattcaattctttaa

Sequence 2130
MIDIPLDEKSFMFDTPGIIQSHQMTNYVYENELKIIIPKNEIKQRVYQLNEKQTLFFGGL
ARIDYVSGGKRPLVCFFSNDLNIHRTKTEKANDLWKSQLGALLSPPQDAQQFNLNDVKAV
RLETGKTKRDIMISGLGFITIDAGAKVIVRVPKHVDVILRNSIL*

Sequence 2131
Contig_0714_pos_6740_7852,
is similar to (with p-value 0.0e+00)
>pir:pir|A55856|A55856 llm protein - Staphylococcus aureus
gtgaggtacaacttattcaatgaaggtgaactgatgtatacactattacttatagctttt
actatgatagtcagtttaataattacacccattattattgtaatatcaaaaaaattagat
ttagtagatcgtcctaatttcagaaaagtacatacgaaacctatctcagtgatgggagga
acggtcattttattttctttcttaatagggatttggctcggacaccctattgaacgtgag
gttaaaccgcttatattaggtgcaattacaatgtatatggttggattgattgatgatatt
tacgatcttaagaccttatttaaagttagcaggtcaaattgttgcagctttaattgttacg
ttttatggaattacaatagactttatttcattgccaattggtccaacgattcattttggc
atattcagcattcctattacagtaatatggattgtagcaattaccaatgctattaatctt
atcgacggacttgatggacttgcctcaggcgtctcagcattggcattaatgactattgga
ttcatcgctattttacaagcgaacatatttattatcatgatttgctgtgtacttttaggg
tctttacttggtttcttattctataactttcacccagcgaaaattttcctaggtgatagt
ggtgcattaatgataggatttattatcggtttcttatccttactcggctttaagaatatc
acatttattgcattattctttcctatagttatattagcggtgccatttattgatacatta
tttgcaatgattcgtcgaatgaaaaaagggcaacatataatgcaagcggacaagtcacat
ttacatcataaattacttgctttaggatatacgcatagacaaaccgttttacttatttat
tcaatagcgattatgtttagtttatctagtgttatcctctatttatcccaaccgttgggt
gcacttatgatgttcattctcattgtctttacgattgagttgatcgttgaatttactgga
ttaatagatgataattatcgaccaatattaaatttaattacaaaaaaaggaaatggtaag
caacatcattatgatgagcatcaccgttcataa

Sequence 2132
VRYNLFNEGELMYTLLLIAFTMIVSLIITPIIIVISKKLDLVDRPNFRKVHTKPISVMGG
TVILFSFLIGIWLGHPIEREVKPLILGAITMYMVGLIDDIYDLRPYLKLAGQIVAALIVT
FYGITIDFISLPIGPTIHFGIFSIPITVIWIVAITNAINLIDGLDGLASGVSALALMTIG

FIAILQANIFIIMICCVLLGSLLGFLFYNFHPAKIFLGDSGALMIGFIIGFLSLLGFKNI
TFIALFFPIVILAVPFIDTLFAMIRRMKKGQHIMQADKSHLHHKLLALGYTHRQTVLLIY
SIAIMFSLSSVILYLSQPLGALMMFILIVFTIELIVEFTGLIDDNYRPILNLITKKGNGK
QHHYDEHHRS*

5

Sequence 2133
Contig_0714_pos_8431_7913,
is similar to (with p-value 7.0e-41)
>gp:gp|Z99122|BSUB0019_48 Bacillus subtilis complete genome
10   (section 19 of 21): from 3597091 to 3809700. NID: g2636029.
>gp:gp|U56901|BSU56901_2 Bacillus subtilis putative transcri
ptional regulator (yvhJ), Ycr59c/YigZ homolog (yvhK), histid
ine kinase (degS),transcriptional regulator of degradation e
nzyme (degU), (degV), (comFA), (comFB), (comFC), flagellar p
15   rotein (yviB), negative regulator of flagellin (flgM), flage
llar protein (yviC), flagellar-hook associated protein 1 (fl
gK), flagellar-hook associated protein 3 (flgL), (yviE), tra
nsmembrane protein (yviF), (csrA), flagellin (hag), flagella
r protein (yviH), flagellar hook-associated protein 2 (fliD)
20   , flagellar protein (fliS), flagellar protein (fliT), sigma-
54 modulator homolog (yviI), and (secA) genes, complete cds.
      NID: g1762326.
      gtgatccgttcgatacataaagatgcaactcataattgttcagcctatactgtcggacca
      gagatgaatattcaaaaggcaaacgacgatggcgaaccaagtggaacagctggcatccca
25   atgcttgaaatactgaaaaaacaagagatacacaatgtttgtgtcgtcgtgacacgctac
      ttcggtggtatcaagttaggtgcaggcggtcttattagagcatatagcggcgcgccgtgcgt
      gatgtgatatatgatatagtagagtcgaactaagagaagctattccagtaaccgttacg
      ttagattatgatcagacaggtaaatttgaatatgaacttgcctctactacattcttatta
      agagaacaattttataccgataaagtaagttatcaaattgacgtagtaaaaaatgaatat
30   gatgcttttatagacttttttaaatcgaattacttctggaaattatgatttgaaacaagaa
      gaccttaaactattaccttttgatattgaaaccaattaa

Sequence 2134
VIRSIHKDATHNCSAYTVGPEMNIQKANDDGEPSGTAGIPMLEILKKQEIHNVCVVVTRY
35   FGGIKLGAGGLIRAYSGAVRDVIYDIGRVELREAIPVTVTLDYDQTGKFEYELASTTFLL
REQFYTDKVSYQIDVVKNEYDAFIDFLNRITSGNYDLKQEDLKLLPFDIETN*

Sequence 2135
Contig_0714_pos_6433_5357,
40   is similar to (with p-value 2.0e-25)
>sp:sp|P54595|YHCK_BACSU HYPOTHETICAL 40.7 KD PROTEIN IN CSP
B-GLPP INTERGENIC REGION. >gp:gp|X96983|BS75DGREG_12 B.subti
lis chromosomal DNA (region 75 degrees: cspB upstream of glp
PFKD operon). NID: g1239975. >gp:gp|Z99108|BSUB0005_180 Baci
45   llus subtilis complete genome (section 5 of 21): from 802821
      to 1011250. NID: g2633055.
      atggaaatgtttgaagctatcatatataacatatctgtcatggtggcaggtatatattta
      tttcataggttacaatattctgaaaataaaagaatgattttttctaaagaatatgtaaca
      gtactaatgacattcgtttctttacttttagcggcatacctatcccatttcaaaacgaa
50   tacctcgtccatttaacatttgtacctcttttgtttttaggacgttataccaacatgata
      tatacactcacggctgcttttatcgtatctttagtcgatgtatttatctttggaaatca
      attatttтtggtattacattaatcgttattgcaggtattgtcagtgcagtgggaccattc
      ttaaagcaaaacgatatcatttctttacttattttaaatttgattagcattatcattttg
      ttatttttagcattattaagccctatttatgaactcgtagagatttagtgcttatccct
55   atttcatttattattacaattgcttcagcaataacattcgttgatatatggcactttttc
      tctttagtcaatcgttatgaaaatgaagataaatacgattatcttacaggtctaggtaat
      gtgaaagaatttgatagacacttaaatgaggtctcaagtaaagctgaagaaaagaaacaa
      agtttagccttacttctcattgatattgatggctttaaagatgtaaacgatcattattca
      caccaatcaggagatgctgttctcaaacaaatgtctcaactattaaaaaactatgtccca

aaccagttcaaaatatttagaaacggtggcgaagaattttctgttgtaataagagattac
acactagatcaaagcgtgaaattagcagaaaatattcgaagtggtgttgaaaaatcttct
ttccacctaccaaacaaagaagtaatcaagctatcagtttcaattggtgtaggatactta
ac'.ccagaagatcgtaaatctcaacgtaaagtatttaaagatgctgatgacatgg'.acat
5   gtggctaaaagtgaaggaagaaataaagtcatgtttaatcctattgtcaaattataa

Sequence 2136
MEMFEAIIYNISVMVAGIYLFHRLQYSENKRMIFSKEYVTVLMTFVSLLLAAYPIPFQNE
YLVHLTFVPLLFLGRYTNMIYTLTAAFIVSLVDVFIFGNSIIYGITLIVIAGIVSAVGPF
10   LKQNDIISLLILNLISIIILLFLALLSPIYELVEILVLIPISFIITIASAITFVDIWHFF
SLVNRYENEDKYDYLTGLGNVKEFDRHLNEVSSKAEEKKQSLALLLIDIDGFKDVNDHYS
HQSGDAVLKQMSQLLKNYVPNQFKIFRNGGEEFSVVIRDYTLDQSVKLAENIRSGVEKSS
FHLPNKEVIKLSVSIGVGYLTQEDRKSQRKVFKDADDMVHVAKSEGRNKVMFNPIVKL*

15   Sequence 2137
Contig_0714_pos_5051_4662,
putative peptide of unknown function
atgaatcgtattgcccatagttatggtttacatgatacatacagttttgtgacatcaact
gcaattattttctcattaaatgatcgtactagtacgaggttgattcgtattcgcgaacgt
20   acaaccgatcttgagaaaattgctttaaccaatagcctatctcgtaaaatttcgagtaag
caacttacaattgacgaagcaaaaagtgagttactgcaacttaaacgtgcgtctcttcag
tattctttcttaacaaatctcattgctgcctttgtagcttgtggttttttcttattcatg
tttggtggcgtagcttccgacgcttggattgcatgcctagcgggtggcatagctttttta
acgt'.tagtttcgtgcactttctatctaa
25

Sequence 2138
MNRIAHSYGLHDTYSFVTSTAIIFSLNDRTSTRLIRIRERTTDLEKIALTNSLSRKISSK
QLTIDEAKSELLQLKRASLQYSFLTNLIAAFVACGFFLFMFGGVASDAWIACLAGGIAFL
TFSFVHFSI*
30

Sequence 2139
Contig_0714_pos_2877_2557,
is similar to (with p-value 5.0e-30)
>sp:sp|P54453|YQEH_BACSU HYPOTHETICAL 41.0 KD PROTEIN IN NUC
35   B-AROD INTERGENIC REGION. >gp:gp|D84432|BACJH642_92 Bacillus
subtilis DNA, 283 Kb region containing skin element. NID: g
2627063. >gp:gp|Z99117|BSUB0014_47 Bacillus subtilis complet
e genome (section 14 of 21): from 2599451 to 2812870. NID: g
2634966.
40   atgatacctggtgtatcaaacataaatgatttttcgtctaatggaatatctatcatatct
aaagttgttcctggaaagcgtgatgtagttactacatcttttttctcccacactctgttca
attaatttattaattaatgtagattttccaacattcgttgtacctacaatgtatacgtca
tctttatt'.cttacatggtttatagattgcaataattcatcaatcccccaaccttt'.attt
gcagaaataagaacgacatcttctgcttctaatccatatttacgagcagattttct'.caac
45   cattcttttacacgtcgatga

Sequence 2140
MIPGVSNINDFSSNGISIISKVVPGKRDVVTTSFSPTLCSINLLINVDFPTFVVPTMYTS
SLFLTWFIDCNNSSIPQPLFAEIRTTSSASNPYLRADFLNHSFTRR*
50

Sequence 2141
Contig_0714_pos_644_213,
is similar to (with p-value 1.0e-21)
>gp:gp|D50453|D50453_106 Bacillus subtilis DNA for 25-36 deg
55   ree region containing the amyE-srfA region, complete cds. NI
D: g1805369.
atgagaattgcacctaatgtgattggtagaatccaaccattaatcgcaccagctattata
agtaaactcaccggtttaccaataaataagaaaacaaaagttgaaattacaataaatgta
ataacgataagattattttttattgagtaacgatttgtgtagtgtttttaaaaatgttgcg

```
cttgtatatgcagaaccaattactgaggacattgctgctgcaaatattactacgccaaaa
atatttttacctataggacctaatgcatgttggaaaactgatgctggtggattttctgaa
ctaagcgtaacgccagttacaacaacacctagtacagctaaaaacaataaggtgcgcatg
acaccagttgttaaaatacctgctacagcagatcgatttacgaaaggaaggtatgactta
```
5   `cctttttatataa`

Sequence 2142
MRIAPNVIGRIQPLIAPAIISKLTGLPINKKTKVEITINVITIRLFLLSNDLCSVFKNVA
LVYAEPITEDIAAANITTPKIFLPIGPNACWKTDAGGFSELSVTPVTTTPSTAKNNKVRM

10   TPVVKIPATADRFTKGRYDLPFI*

Sequence 2143
Contig_0715_pos_1038_1508,
is similar to (with p-value 4.0e-35)

15   >sp:sp|Q02134|HIS7_LACLA IMIDAZOLEGLYCEROL-PHOSPHATE DEHYDRA
TASE (EC 4.2.1.19) (IGPD). >pir:pir|G45734|G45734 HisB · Lac
tococcus lactis subsp. lactis >gp:gp|U92974|LLU92974_6 Lacto
coccus lactis unknown gene, partial cds, and HisC (hisC), un
known, HisG (hisG), unknown, HisB (hisB), unknown, HisH (his

20   h), HisA (hisA), HisF (hisF), HisIE (hisIE), unknown, unknow
n, LeuA (leuA), LeuB (leuB), LeuC (leuC), LeuD (leuD), unkno
wn, IlvD (ilvD), IlvB (ilvB), IlvN, IlvC (ilvC), IlvA (ilvA)
, AldB (aldB) and aldR (aldR) genes, complete cds. NID: g256
5137.

25   `atgttaacgctatttacttttcatagtggattaactttatctattgaggccactggagat`
```
acgtatgttgatgatcatcatataactgaagatataggtatagttattggacaattactt
cttgaattaataaagactcaacaaagtttttacaagatatggttgctcatatgtacccatg
gatgaggcgcttgctcgaacagtagtggacattagtggtcgtccatatttctcatttaat
agcaagttgagcgctcaaaaggtaggaacttttgacactgaactagttgaagaatttttt
```
30   `agagcattgataattaatgcgcgattaaccgttcacattgacttattaagaggtggaaat`
```
acacatcatgagattgaggcaatatttaaatcttttgcaagagcattaaagatttctctt
gcacaaaatgaagatggacgtattccatcgtctaaaggagtaattgaatga
```

Sequence 2144
35   MLTLFTFHSGLTLSIEATGDTYVDDHHITEDIGIVIGQLLLELIKTQQSFTRYGCSYVPM
DEALARTVVDISGRPYFSFNSKLSAQKVGTFDTELVEEFFRALIINARLTVHIDLLRGGN
THHEIEAIFKSFARALKISLAQNEDGRIPSSKGVIE*

Sequence 2145
40   Contig_0715_pos_1607_1990,
is similar to (with p-value 1.0e-21)
>sp:sp|Q02132|HIS5_LACLA AMIDOTRANSFERASE HISH (EC 2.4.2.-).
 >pir:pir|I45734|I45734 HisH - Lactococcus lactis subsp. lac
tis >gp:gp|U92974|LLU92974_8 Lactococcus lactis unknown gene

45   , partial cds, and HisC (hisC), unknown, HisG (hisG), unknow
n, HisB (hisB), unknown, HisH (hish), HisA (hisA), HisF (his
F), HisIE (hisIE), unknown, unknown, LeuA (leuA), LeuB (leuB
), LeuC (leuC), LeuD (leuD), unknown, IlvD (ilvD), IlvB (ilv
B), IlvN, IlvC (ilvC), IlvA (ilvA), AldB (aldB) and aldR (al

50   dR) genes, complete cds. NID: g2565137.
```
gtgcaaaaagctgaagctatcgtacttccaggtgttggacattttcaggatgcgatgcat
tctatagaagaaaaaagcattaaagatatgcttaaaaatatacatgataaaccgataatt
ggaatatgtttaggtatgcaattacttttttcaacatagcgcagaaggtgacgttagtgga
ttggaacttgtcccgggaaatatagtgccaatccaatcatctcatcctattcctcacttg
```
55·   `ggttggaatgaattaaagagtacacatcccttactgcaaagtgatgtgtattttgttcat`
```
tcatatcaagcagaaatgtcagaatatgtagtagcttatgctgactatggtacaaagatt
ccgggagtcattcaataccgatga
```

Sequence 2146

VQKAEAIVLPGVGHFQDAMHSIEEKSIKDMLKNIHDKPIIGICLGMQLLFQHSAEGDVSG
LELVPGNIVPIQSSHPIPHLGWNELKSTHPLLQSDVYFVHSYQAEMSEYVVAYADYGTKI
PGVIQYR*

5   Sequence 2147
    Contig_0715_pos_2067_2615,
    is similar to (with p-value 9.0e-31)
    >sp:sp|P54471|YQFN_BACSU HYPOTHETICAL 23.7 KD PROTEIN IN CCC
    A-SODA INTERGENIC REGION. >gp:gp|D84432|BACJH642_139 Bacillu
10 s subtilis DNA, 283 Kb region containing skin element. NID:
    g2627063. >gp:gp|Z99116|BSUB0013_228 Bacillus subtilis compl
    ete genome (section 13 of 21): from 2395261 to 2613730. NID:
     g2634723.
    gtgattcaaggaccttataaggctgctaaaagaaatattgcaaattatgaattaaatcaa
15 caggttgatgtacgtctaggcgatggtctaagcgttataaacacagaagaccaaattgat
    aatataactgtttgtggtatgggagggccattaattgcaaaaatattaaacgatggaaaa
    gataaattagttaaccatccaagactcatactacaaagcaacatacaaactcaagcatta
    agacaaactcttaataaactttcatatgaaatcgttgatgaaagaatcattgaggaaaag
    ggtcacatatatgaaatcgtggtagctgagtttaataataacttagttaaattaaatata
20 ttacaagaaaaattcggaccatttttacttagagaatgtaataacattttttcaaaaaaaa
    tggcaaagagagttagaagcactgcgtgatataaaatcccaattgaattcaacatcacat
    catgagagactaaaagaaatagaagatgaaattaacttaatacaagaggtgttaattaat
    gaaaattag

25 Sequence 2148
    VIQGPYKAAKRNIANYELNQQVDVRLGDGLSVINTEDQIDNITVCGMGGPLIAKILNDGK
    DKLVFHPKLILQSNIQTQALRQTLNKLSYEIVDERIIEEKGHIYEIVVAEFNNNLYKLNI
    LQEKFGPFLLRECNNIFQKKWQRELEALRDIKSQLNSTSHHERLKEIEDEINLIQEVLIN
    EN*
30
    Sequence 2149
    Contig_0715_pos_2623_3705,
    is similar to (with p-value.2.0e-54)
    >sp:sp|P53434|YRP2_LISMO HYPOTHETICAL 41.4 KD PROTEIN IN RPO
35 D 3'REGION (ORFA2). >gp:gp|U17284|LMU17284_3 Listeria monocy
    togenes major sigma factor (rpoD) gene, partial cds, and dow
    nstream orfA1 and orfA2 genes, complete cds. NID: g687597.
    atggaagttttaaataatcacgttccatttcatcaagctgaatcatgggataatgtagga
    ttattaattggtaatgataagttagatattacaggtattctgacaacactcgactgcacc
40 gatgatgttgttaaccaagcaattgaacttaataccaataccatcattgctcatcatcca
    cttattttcaaaggagtaaaacgtatcgttgaagatggatatggtagtataattcgtaaa
    cttatccaaaataatatcaatcttatagcattacacactaatcttgatgtaaatcctaaa
    ggtgtcaatcgaatgttagcggatcaaataggtttagagaacatatcaatgattaataca
    aatagctcatattattacaaagttcaaactttatacctaaaaattatattgaagatttc
45 aaagacagtttaaacgaacttggattagctaaagaaggtaattacgaatattgtttcttt
    gaaagtgaaggtaaagggcaatttaaaccagtaggtgatgcaagtccttatataggyaag
    ttagatagtatcgaatatgttgatgaaataaaacttgagtttatgataaaagacaatgaa
    ttagaaataactaaacgtgctattttagataatcacccatacgaaacaccagtttttgat
    tttattaaaatgaacaaagaaagtgagtatggattagggattattggacaattaaaccaa
50 actatgactttagatgaattttctgaatatgccaaaaaacagctcaatataccgagcgta
    cgatatacaggtcaacatgatagtccaattaagaaagtagctatcataggtggttcaggt
    ataggatttgagtataaagctagccaacttggagcagatgttttgttactggtgatatt
    aaacaccatgatgctttagatgctaaaatccaaatgtaaatttattagacatcaatcat
    tatagtgagtatgttatgaaagaaggattaaaagaattattagaaaaatggttatttaaa
55 tatgaaaatcaatttccaatatatgcttctgaaatcaacacagatccatttaaatataaa
    taa

    Sequence 2150
    MEVLNNHVPFHQAESWDNVGLLIGNDKLDITGILTTLDCTDDVVNQAIELNTNTIIAHHP

LIFKGVKRIVEDGYGSIIRKLIQNNINLIALHTNLDVNPKGVNRMLADQIGLENISMINT
NSSYYYKVQTFIPKNYIEDFKDSLNELGLAKEGNYEYCFFESEGKGQFKPVGDASPYIGK
LDSIEYVDEIKLEFMIKDNELEITKRAILDNHPYETPVFDFIKMNKESEYGLGIIGQLNQ
TMTLDEFSEYAKKQLNIPSVRYTGQHDSPIKKVAIIGGSGIGFEYKASQLGADVFVTGDI
KHHDALDAKIQNVNLLDINHYSEYVMKEGLKELLEKWLFKYENQFPIYASEINTDPFKYK
*

Sequence 2151
Contig_0715_pos_3740_4072,
putative peptide of unknown function
atgtcaaaacatccatttgaacactttaatttagatgagaatttaattgaagctgttaaa
aatctcaattttgaaaaaccgactgaaatccaaaatagaatcataccgagaattcttaaa
ggaacaaatttaataggacaatctcaaactggaactggaaagtcacacgcttttctttta
atattgatggagatcgttctttttctgaagcagtacaattggacatacaagtacctggtg
cttttcctaaatttggaactacagatggaacgattacgaaaaatgttgtcattgaaaaat
gttattttggctgttcagatcatcctaaaatga

Sequence 2152
MSKHPFEHFNLDENLIEAVKNLNFEKPTEIQNRIIPRILKGTNLIGQSQTGTGKSHAFLL
ILMEIVLFLKQYNWTYKYLVLFLNLELQMERLRKMLSLKNVILAVQIILK*

Sequence 2153
Contig_0717_pos_4083_5096,
putative peptide of unknown function
atggaacgattttgttgtgtaaatcaaattaactatattcaaatgaatccgttagaagcc
aaatttaaaacgagcgctctaagatcatggaaaactgatcaggcagatgctcataagctt
gcttgtttaggaccgacgcttaaacaaacagacaacttacctatacatgagttaatattc
tttgaattaagagaacgcgtccgttttcatctagaaatcgagaatgaacaaaatcgactt
aaatttcagatccttgaattactccatcaaacattccctggtttagaaagattatttagt
agtcgatattcaatcattgcactcaacatcgcagaaatctttactcatccagacatggtt
cttgatatcgacaaggaggtactgattacacatatattcaattctacagataagggaatg
tcaatggataaagctacaaaatatgcacttcaattaagggtgattgctcaagaaagctat
cctaatgtcgatagacattcctttctagtcgaaaaattacgcttacttattcaacaatta
aaacaatctattcatcatatcaaacaattagatgatgccatgattcaattagcacaacaa
ctcgattattttgaaaatattcattcgatacctggtattggtaagctaagcacagctatg
attattgggaagattggtgatattaagcgatttaaatcaaataaacaactcaatgctttt
gttggcattgatatcaaacgatatcaatcaggtcatacacactgtagagataccatcaac
aagcctgtgtaataaaaaagcgagaaaactttattttgggtgattatgaatataataaga
gggcagcatcattatgacaatcatgtcgtcgattattactacaaactaagaaagcagcct
aatgagaaacctcataagactgccatcattgcttgtataaatcgattattaaaaacaatt
cattatcttgtaatgaattataaattgtacgattatcaaatgtcaccacattag

Sequence 2154
MERFCCVNQINYIQMNPLEAKFKTSALRSWKTDQADAHKLACLGPTLKQTDNLPIHELIF
FELRERVRFHLEIENEQNRLKFQILELLHQTFPGLERLFSSRYSIIALNIAEIFTHPDMV
LDIDKEVLITHIFNSTDKGMSMDKATKYALQLRVIAQESYPNVDRHSFLVEKLRLLIQQL
KQSIHHIKQLDDAMIQLAQQLDYFENIHSIPGIGKLSTAMIIGKIGDIKRFKSNKQLNAF
VGIDIKRYQSGHTHCRDTINKRGNKKARKLLFWVIMNIIRGQHHYDNHVVDYYYKLRKQP
NEKPHKTAIIACINRLLKTIHYLVMNYKLYDYQMSPH*

Sequence 2155
Contig_0717_pos_8693_0,
is similar to (with p-value 1.0e-49)
>sp:sp|P42423|YXDL_BACSU HYPOTHETICAL ABC TRANSPORTER ATP-BI
NDING PROTEIN IN IDH 3'REGION. >gp:gp|D14399|BACIOLO_13 Baci
llus subtilis 15 kb chromosome segment contains the iol oper
on. NID: g709980. >gp:gp|Z99124|BSUB0021_68 Bacillus subtili
s complete genome (section 21 of 21): from 3999281 to 421481
4. NID: g2636442. >gp:gp|D45912|D45912_2 Bacillus subtilis g

enome sequence between the iol and hut operon, partial and c
omplete cds. NID: g1408482.
atgggtccttctggatcaggtaaaacgactttactcaatgtgttaagttcaatagatact
atttcagaaggaactgtggaagttgaaggcaaagaaattaataaactgagccacaaagaa

5   gtggcaaattttcgaaaacaacatctcggttttatttttcaagattatagcgttttaccc
acattaacagtaaaagaaaatattatgctaccactctcagtacaaaaattccataaaatat
gaaatggaacaaaattataaagaagtggctgaggcattaggtatttataacctgggaaat
aaatatccaagtgaaatttctggcggtcagcaacaacgtacggcggcagcccgggcattc
gtccataaaccaacgattattttcgcagatgaacctactggcgcattagattctaaaagt

10   gctcaagatttgttacaccgtctagaagatatgaataaacaatttaattcaaccattatg
atggtgacacatgatccttcagccgctagttacgctgagagagtcattatgttgaaagac
ggtgatatacactcagaaatctaccagggtaacgattcaaaacaaacattttaccaagaa
attatgaaacttcc

15   Sequence 2156
MGPSGSGKTTLLNVLSSIDTISEGTVEVEGKEINKLSHKEVANFRKQHLGFIFQDYSVLP
TLTVKENIMLPLSVQKFHKYEMEQNYKEVAEALGIYNLGNKYPSEISGGQQQRTAAARAF
VHKPTIIFADEPTGALDSKSAQDLLHRLEDMNKQFNSTIMMVTHDPSAASYAERVIMLKD
GDIHSEIYQGNDSKQTFYQEIMKLP

20   Sequence 2157
Contig_0717_pos_8309_8007,
putative peptide of unknown function
atggagaaaaggagtattaatatgaaaaaagtatttatgatcataagtatacttaccata

25   actgttactttaagtgcatgtggaggttctggaaaacaaaaagagccatctaaggaaagt
caaaaatctgataaatatgattatgtttattatgaaatattaaatgatggagattctgaa
acgccaaatgttgagattaaatataaagataaaaaaggtaaatcacatatagaaaaagct
gatttagatcacgtgtatgaacatatactaggtgatggtaataaaaaaccatatatgata
tga

30   Sequence 2158
MEKRSINMKKVFMIISILTITVTLSACGGSGKQKEPSKESQKSDKYDYVYYEILNDGDSE
TPNVEIKYKDKKGKSHIEKADLDHVYEHILGDGNKKPYMI*

35   Sequence 2159
Contig_0717_pos_3267_2203,
is similar to (with p-value 0.0e+00)
>sp:sp|P38021|OAT_BACSU ORNITHINE AMINOTRANSFERASE (EC 2.6.1
.13) (ORNITHINE--OXO-ACID AMINOTRANSFERASE). >pir:pir|B53370

40 |B53370 orthinine aminotransferase - Bacillus subtilis >gp:g
p|D78193|BACGNTZA_27 Bacillus subtilis 36kb sequence between
 gntZ and trnY genes encoding 34 ORFs. NID: g1064780. >gp:gp
|X81802|BSROCDEF_1 B.subtilis rocD, rocE and rocF genes. NID
: g550310. >gp:gp|Z99124|BSUB0021_139 Bacillus subtilis comp

45   lete genome (section 21 of 21): from 3999281 to 4214814. NID
: g2636442.
atggatatgctttcggcctactcggcagtgaatcaaggtcatcgacacccaagaattatt
caagcattgaaagatcaagcagataaagtcactttagtatcacgtgctttcatagtgat
aatccgggtcaatggtatgagaaaatatgtaaactcgcaggtaaagacaaagcattgcct

50   atgaatacgggagcagaggcggttgaaacagctttaaaagctgctcgtcgttgggcttat
gatgttaagggtattgagccgaacaaagctgaaattatcgcttttaatggtaatttccat
ggacgtacgatggcaccagtatcattgtcttcagaagctgagtatcaacgaggctatggt
ccattgttagatggctttcgaaaagttgagtttggtgacgttaatcaattaaaagcagca
attaataaaaatacagcagcaatttttagtagaacctatacagggagaagcagggattaac

55   gtaccaccagaaggatatttgaaaacaattagagaattatgtgatgaacatcaaatttta
tttattgctgatgaaattcaagcaggattaggacgttcaggaaaattatttgcaacggat
tgggatcatgtaaaaccggatgtttatattttaggaaaagcgttaggtggaggggtattt
cctatctcggtagttcttgcagataatgaggtattagatgtatttactcctggctcacat
ggttctacatttggtggaaatccactagcgagtgcagttctattgcagctatagatgtc

536

attaːtgːacgaggatttacctggtcgttcattagaattaggagaatattttaagtːtgaa
ttgaaaaaaattgagcatccatctattaaagaagttaggggacgaggattatttaːcggt
attgaattacatgaaagtgccagaccatattgtgaagctttgaaagaacaaggattatta
tgtaaagaaactcacgacaccgttattagatttgcacctccattagtgataacgaaagaa
5   gagttagacatggctttagagaagattaagagtgtatttgcatag

Sequence 2160
MDMLSAYSAVNQGHRHPRIIQALKDQADKVTLVSRAFHSDNLGQWYEKICKLAGKDKALP
MNTGAEAVETALKAARRWAYDVKGIEPNKAEIIAFNGNFHGRTMAPVSLSSEAEYQRGYG
10   PLLDGFRKVEFGDVNQLKAAINKNTAAILVEPIQGEAGINVPPEGYLKTIRELCDEHQIL
FIADEIQAGLGRSGKLFATDWDHVKPDVYILGKALGGGVFPISVVLADNEVLDVFTPGSH
GSTFGGNPLASAVSIAAIDVIIDEDLPGRSLELGEYFKSELKKIEHPSIKEVRGRGLFIG
IELHESARPYCEALKEQGLLCKETHDTVIRFAPPLVITKEELDMALEKIKSVFA*

15   Sequence 2161
Contig_0717_pos_1974_850,
is similar to (with p-value 0.0e+00)
>sp:sp|P50735|YPCA_BACSU HYPOTHETICAL 46.7 KD OXIDOREDUCTASE
     IN RECQ-CMK INTERGENIC REGION (EC 1.4.1.-). >gp:gp|Z99115|B
20   SUB0012_236 Bacillus subtilis complete genome (section 12 of
     21): from 2195541 to 2409220. NID: g2634478. >gp:gp|Z99116|
     BSUB0013_8 Bacillus subtilis complete genome (section 13 of
     21): from 2395261 to 2613730. NID: g2634723. >gp:gp|L47648|B
     ACSERA_11 Bacillus subtilis phosphoglycerate dehydrogenase (
25   serA), ypaA, ferredoxin (fer), ypbB, recS, ypbD, ypbE, ypbF,
     ypbG, ypbH, glutamate dehydrogenase (ypcA), ypdA, ypdB, ypd
     C, spore cortex lytic enzyme (sleB), ypeB, ypfA, ypfB, cytid
     ine monophosphate kinase (cmk), ypfD, ypgA, yphA, yphB, yphC
     , NAD+ dependent glycerol-3-phosphate dehydrogenase (glyc),
30   yphE and yphF genes, complete cds. NID: g1146195. >gp:gp|L47
     648|BACSERA_11 Bacillus subtilis phosphoglycerate dehydrogen
     ase (serA), ypaA, ferredoxin (fer), ypbB, recS, ypbD, ypbE,
     ypbF, ypbG, ypbH, glutamate dehydrogenase (ypcA), ypdA, ypdB
     , ypdC, spore cortex lytic enzyme (sleB), ypeB, ypfA, ypfB,
35   cytidine monophosphate kinase (cmk), ypfD, ypgA, yphA, yphB,
     yphC, NAD+ dependent glycerol-3-phosphate dehydrogenase (gl
     yc), yphE and yphF genes, complete cds. NID: g1146195.
gtgcgcattccagtacgaatggacgatgggacagttaaaacatttactggataccgtgca
caacataatgatgcagttggccctactaaaggtggcgttcgtttccatccagaagttgat
40   gaagaagaagtgaaggcattatcaatgtggatgactttaaaatgtggcattgttaːːtta
ccttatcgtggggggtaaaggcggtattgtttgtgatccacgtcaaatgagtatacːcgaa
gttgagcgtctatctcgagggtatgtaagagctatatctcaatttgttgggcctaacaaa
gatattccagcaccagacgttttacaaactctcaaattatggcttggatgatggatgag
tatagtgcattagataaatttaattcgccaggttttataacaggcaaacctattgtatta
45   ggtggatctcaaggacgtgatcgctctacagctttaggtgtagtcattgcaattgaacaa
gcggcaaaacgtcgtggcatggatattaaagatgccaaaattgtgattcaaggtttcggc
aatgcaggtagtttcttagctaaattcttatacgatttaggtgctaaagtagttggtata
agtgatgcatatggagccttacatgaccctaatggacttgatatagattatttattagat
cgacgtgatagcttcggtacagttacaaatctattcgaagatacaatttctaacaaagaa
50   ttattcgaattggattgtgacatccttgttcctgctgcaatttccaatcaaatcactgaa
gataatgcgcacgatattaaagcaagcattgtagttgaagctgctaatggccctacgacg
cctgaagcaacacgtatttttaacagaaagagatatattactagtgccggatgtacttgca
agtgcaggaggtgtgactgtatcttatttcgagtgggtacaaaataatcaaggttattat
tggactgaagaagaagtcaatgacaaactacgtgagaagttagtaacagcatttgatacg
55   atttatgaattgtcacaaaatagaaaaattgatatgagattagcagcatatatagtaggt
attaaacgtactgctgaagcagcaagatatagaggttgggcataa

Sequence 2162
VRIPVRMDDGTVKTFTGYRAQHNDAVGPTKGGVRFHPEVDEEEVKALSMWMTLKCGIVNL

PYGGGKGGIVCDPRQMSIHEVERLSRGYVRAISQFVGPNKDIPAPDVFTNSQIMAWMMDE
YSALDKFNSPGFITGKPIVLGGSQGRDRSTALGVVIAIEQAAKRRGMDIKDAKIVIQGFG
NAGSFLAKFLYDLGAKVVGISDAYGALHDPNGLDIDYLLDRRDSFGTVTNLFEDTISNKE
LFELDCDILVPAAISNQITEDNAHDIKASIVVEAANGPTTPEATRILTERDILLVPDVLA
5   SAGGVTVSYFEWVQNNQGYYWTEEEVNDKLREKLVTAFDTIYELSQNRKIDMRLAAYIVG
IKRTAEAARYRGWA*

Sequence 2163
Contig_0718_pos_4955_5407,
10   putative peptide of unknown function
atgtatgatgaattagaggttaataaaagtcgttttaaaaactgtaattttaatgaaggt
attttttaagaatatagaagcaatttgtaattgtaaatttacaacgtgcgggtttaataat
tgtattttcgaagatgttcatttttacaaaaaccaatttaaagattcaacatttgtgaat
acaccatttgatcaatccgtatttaatagtactttattccaaaatgcaatgttcgatagc
15   aatctcattcgtagcgtaaaatggactgatatcattttaaaaacgtttctttcaaaaat
gtagaaattgaaggaacaacatttaaagatgtaaaattcaaaaattgtgagttcaaaaat
gtaattattactaattcaactatgtcgcaaaagttaatgaatgaattacaaaaacaagat
gttactttagaaaatatagacacttctatttaa

20   Sequence 2164
MYDELEVNKSRFKNCNFNEGIFKNIEAICNCKFTTCGFNNCIFEDVHFYKNQFKDSTFVN
TPFDQSVFNSTLFQNAMFDSNLIRSVKWTDIIFKNVSFKNVEIEGTTFKDVKFKNCEFKN
VIITNSTMSQKLMNELQKQDVTLENIDTSI*

25   Sequence 2165
Contig_0718_pos_8056_7316,
is similar to (with p-value 3.0e-34)
>sp:sp|Q06174|EST_BACST CARBOXYLESTERASE PRECURSOR (EC 3.1.1
.1). >pir:pir|JC1374|JC1374 carboxylesterase (EC 3.1.1.1) −
30   Bacillus stearothermophilus (strain IFO 12550) >gp:gp|D12681
|BACPBH7_1 Bacillus stearothermophilus esterase gene. NID: g
216313..
gtgaatgtaaaaatgaaagttaaatcaccacaatcaatctacttaaaaggacatcgtcaa
caagctgtattgttattacattcttttacgggaactgtacgtgatgtaaaacatttagca
35   caacagttgaatgaagagggatttacttgttacgtgcctagttatccaggccacggtttg
ccacttaaggaatttacccaacacaatattaatgattggtgggaacaagttacagcagca
tatcaattttttaagaaatgcaggatacagtagaattaatgtgacaggcgtatcattaggc
ggattatttactttaaggttagctgaacattttgatttagaacgtatagctgtgatgtca
gccccacataaaaagcgtgaaagcgagattgcgtggcgtcttgaaaggtatgggcatcga
40   atgaatgaaatttttgagtttaagcgaagaagagcgtcgtcaccaaatggaaaccatcttg
tcttatgataaagaaattgaagtgtttcaaggtgtaattgatgaaattatggcttatctt
gcaaatattacagtaccagtgaatattatgtatggcgaagaagatgacccattatatgct
caaagtgcgcaatacatttatgataatgtaaatagtcaagataaagaactgctcaaattt
gaaaaaagcggtcatcttatgacgtatggcgatcatgcatacagagtagaacaatctatt
45   attcaatttttcagtaaataa

Sequence 2166
VNVKMKVKSPQSIYLKGHRQQAVLLLHSFTGTVRDVKHLAQQLNEEGFTCYVPSYPGHGL
PLKEFTQHNINDWWEQVTAAYQFLRNAGYSRINVTGVSLGGLFTLRLAEHFDLERIAVMS
50   APHKKRESEIAWRLERYGHRMNEILSLSEEERRHQMETILSYDKEIEVFQGVIDEIMAYL
ANJTVPVNIMYGEEDDPLYAQSAQYIYDNVNSQDKELLKFEKSGHLMTYGDHAYRVEQSI
IQFFEK*

Sequence 2167
55   Contig_0718_pos_6737_5586,
putative peptide of unknown function
atggaatttacggtaggtaaaatgggtcgtacatatacaacgcaaatatataagaaatta
acgggaaagaaatggcttaatattatcggatggaatggtaatttagccgtatttatacta
tttggttttttatagtgttattggtggttggattattatatatataggttatgtcatagca

```
caaatcatggttttaaatcaagtacgctgacaaatattcaatttgaaacaatcattagt
aatccatggttgactgttttaggtcaaggcatatttattttgataacaatggtaattgtt
atgttaggtgttgaaaaaggtttagaaaaagcttctaaaataatgatgcctctattattt
atctctt·.aattatcgttgtagcacaatctttaactttagaaggtgctttagaag.tgta
cgttatatactgcaacctcgagttgaagatatgtctattcaaggtgtactatttgcgtta
ggacaatcgtttttacgctgtccctaggtacaaccggaatgattacttatgcaagctat
gcacctaaaaatatgacgataaagtcttcagcactttcaattgtcgtaatgaatattta
atttctgtcttggctggattagctatatttcctgcgcttaaaacatttggttaccaaccc
caagaaggccctggcttattatttaaggttttaccactagtatttagcgaaatgactttt
ggtacattcttttactttatatttttactattattcttatttgcggcattaacgtcttct
atatcattattagagttaaatgtatctaattttactaaaaatgataatagtaaaagacaa
aaagtggcaatcataggtagtatacttgtatttatcattagtatcccagcaacattatct
tttagtagtctaagtcatttgcgttttggcgctggtacgatatttgataatatggatttt
attgtatctaatattcttatgccattaggggcactaggaacaacattagtggttggccaa
ttactagataaaaaattattaaaagaaagctttgggaaagacaaattcaacctattttta
ccgtggtattatttaattaagttcatcatgcctattgttattattttagtatttatagtt
caattattttaa
```

Sequence 2168

```
MEFTVGKMGRTYTTQIYKKLTGKKWLNIIGWNGNLAVFILFGFYSVIGGWIIIYIGYVIA
QIMVFKSSTLTNIQFETIISNPWLTVLGQGIFILITMVIVMLGVEKGLEKASKIMMPLLF
IFLIIVVAQSLTLEGALEGVRYILQPRVEDMSIQGVLFALGQSFFTLSLGTTGMITYASY
APKNMTIKSSALSIVVMNILISVLAGLAIFPALKTFGYQPQEGPGLLFKVLPLVFSEMTF
GTFFYFIFLLLFLFAALTSSISLLELNVSNFTKNDNSKRQKVAIIGSILVFIISI?ATLS
FSSLSHLRFGAGTIFDNMDFIVSNILMPLGALGTTLVVGQLLDKKLLKESFGKDKFNLFL
PWYYLIKFIMPIVIILVFIVQLF*
```

Sequence 2169
Contig_0718_pos_4565_3657,
is similar to (with p-value 9.0e-76)
>gp:gp|U93874|BSU93874_1 Bacillus subtilis cysteine synthase
 (yrhA), cystathionine gamma-lyase (yrhB), YrhC (yrhC), YrhD
 (yrhD), formate dehydrogenase chain A (yrhE), YrhF (yrhF),
 formate dehydrogenase (yrhG), YrhH (yrhH), regulatory protei
 n (yrhI), cytochrome P450 102 (yrhJ), YrhK (yrhK), hypotheti
 cal protein YrhL (yrhL), putative anti-SigV factor (yrhM), R
 NA polymerase sigma factor SigV (sigV) and YrhO (yrhO) genes
 , complete cds, and YrhP (yrhP) gene, partial cds. NID: g193
 4604. >gp:gp|Z99117|BSUB0014_206 Bacillus subtilis complete
 genome (section 14 of 21): from 2599451 to 2812870. NID: g26
 34966.

```
atgattgcatacgatttgataggacaaactccattagttttattagaaagctttagtgac
gaga.tgt.taaaatatacgccaaacttgagcaatttaatcctggtggtagcatcaaagac
cgtctagggaagtacttaattgaaaaagcaatagatgaaggacgacttaaagaag;ggat
acaatagtcgaagcgactgctggtaatacaggcattggacttgctattgcttctaatcgg
cacaaagtaaaatgtatcatctttgctccagaaggatttgcagaagaaaaaatttcaatt
atgaaagcatttgggtgcagatgttagacgtacccccaaagctgagggaatgactggcgca
cagcaagaggcgttggcatacgcaacacgatatggatatttatatatgaatcaattcgaa
actaaagataatcctggggcatatacacaaacacttgccaaacaactcacagatgaactt
tcacatattgattattttgtggcaggtgttgggtccggtggtacgtttacaggagttgca
caacacttaaaaacgtatgatgtaaaaaattatattgtagaaccagaaggctctgtctta
aatggtggtgtcagtcatcctcatgcaactgaagggattggttctgaaaagtggccatca
tttttagaaaaagaattagtagatggtattttttactgttgccgataaagatgctttttaat
aatgttaaacttgtcgcgaataaagaaggattgttagttggtagttcttcgggagcggca
ttacaaggagcgttggaattaaaaaaaagcattcaaaatggtgtgattgttaccatctttt
ccagatggaagcgatcgatacatgtccaaacaaatattcaactataaggagagttttaat
aatgaataa
```

Sequence 2170

MIAYDLIGQTPLVLLESFSDENVKIYAKLEQFNPGGSIKDRLGKYLIEKAIDEGRLKEGD
TIVEATAGNTGIGLAIASNRHKVKCIIFAPEGFAEEKISIMKALGADVRRTPKAEGMTGA
QQEALAYATRYGYLYMNQFETKDNPGAYTQTLAKQLTDELSHIDYFVAGVGSGGTFTGVA
QHLKTYDVKNYIVEPEGSVLNGGVSHPHATEGIGSEKWPSFLEKELVDGIFTVADKDAFN
5   NVKLVANKEGLLVGSSSGAALQGALELKKSIQNGVIVTIFPDGSDRYMSKQIFNYKESFN
NE*

Sequence 2171
Contig_0718_pos_3646_2519,
10   is similar to (with p-value 0.0e+00)
>gp:gp|U93874|BSU93874_2 Bacillus subtilis cysteine synthase
(yrhA), cystathionine gamma-lyase (yrhB), YrhC (yrhC), YrhD
(yrhD), formate dehydrogenase chain A (yrhE), YrhF (yrhF),
formate dehydrogenase (yrhG), YrhH (yrhH), regulatory protei
15   n (yrhI), cytochrome P450 102 (yrhJ), YrhK (yrhK), hypo'·neti
cal protein YrhL (yrhL), putative anti-SigV factor (yrhM), R
NA polymerase sigma factor SigV (sigV) and YrhO (yrhO) genes
, complete cds, and YrhP (yrhP) gene, partial cds. NID: g193
4604. >gp:gp|Z99117|BSUB0014_205 Bacillus subtilis complete
20   genome (section 14 of 21): from 2599451 to 2812870. NID: g26
34966.
atgatacatgggggacatacgacagacaactatactggagcagtgacaacacctatttat
caaacaagtacttatttacaagatgatattggtgatttaagacaagggtacgaatattca
cgtactgcaaatcctacacgtgcgtctcttgaaagtgttattgctaatttagaacatggt
25   aagcatggttttgcttttggttcaggaatggcagcaattagtgcagttatcatgttatta
gataaaggagatcacttagttcttaattctgatgtttatggtggcacatatcgtgcatta
actaaagtatttactcgctttggtatagacgtagattttgttgatacaactaaaattgaa
aacattgaacaatatattaaacctgaaactaaaatgttatatgtagaaacaccttcaaat
ccattattgcgtgtgactgatattaaagcatcagcaaaaattgcaaaaaaatatgatttg
30   atatctgtagtcgataatacatttatgacaccttactaccaaaaccctttagactttggt
attgatatcgtattgcattcggctacaaaatatattggaggccatagtgatgttgtagct
ggtcttgttgctactgctgatgatgatttagcagaacgtctaggctttatttcaaattct
acaggtggtgtacttggacctcaagatagctatttattaatcagaggtattaaaacgcta
ggtctaagaatggagcaaataaaccgaaacgttgaaggtattgtgcaaatgttacaaaag
35   caccctaaagttcaacaagtattccatcctagtattaaggaacatatgaactatactatc
catcaaaatcaagcaactgggcatacaggggtagtatcttttgaagttaaagatacagaa
gcggctaaacaagtgattcacgcaacaaactactttacactggcagagagtttaggggca
gttgaaagtctaatttctgtaccggcacttatgacgcatgcgtccatcccatcagatgta
agagccaaggaaggtattacggatggtctcattcgtttatctattggtattgaagacaca
40   gaagacttagttaatgatttagaacaagccttaaatactttgagataa

Sequence 2172
MIHGGHTTDNYTGAVTTPIYQTSTYLQDDIGDLRQGYEYSRTANPTRASLESVIANLEHG
KHGFAFGSGMAAISAVIMLLDKGDHLVLNSDVYGGTYRALTKVFTRFGIDVDFVDTTKIE
45   NIEQYIKPETKMLYVETPSNPLLRVTDIKASAKIAKKYDLISVVDNTFMTPYYQNPLDFG
IDIVLHSATKYIGGHSDVVAGLVATADDDLAERLGFISNSTGGVLGPQDSYLLIRGIKTL
GLRMEQINRNVEGIVQMLQKHPKVQQVFHPSIKEHMNYTIHQNQATGHTGVVSFEVKDTE
AAKQVIHATNYFTLAESLGAVESLISVPALMTHASIPSDVRAKEGITDGLIRLSIGIEDT
EDLVNDLEQALNTLR*
50
Sequence 2173
Contig_0719_pos_3554_2802,
is similar to (with p-value 1.0e-71)
>sp:sp|P29928|SUMT_BACME UROPORPHYRIN-III C-METHYLTRANSFERAS
55   E (EC 2.1.1.107) (UROGEN III METHYLASE) (SUMT) (UROPORPHYRIN
OGEN III METHYLASE) (UROM). >pir:pir|A42479|A42479 S-adenosy
l-L-methionine uroporphyrinogen III methyltransferase - Baci
llus megaterium >gp:gp|M62881|BACCOBA_1 Bacillus megaterium
S-adenosy-L-methionine:uroporphyrinogen III methyltransferas

e (COBA) gene, complete cds. NID: g142694.
gtggttcatatgggggaaagtatatttagttggagctggacctggtgatccagaattaata
acgttaaaaggtttaaaagccattaaagaagccgatgtcatcctttatgaccgacttgta
aataaagaaatacttaattatgcttctccttctactaagttcttctattgcggtaaggat
5   cctcacaggcactccttaccgcaggaagaaacaaataaaatgatggtaaccttagccaaa
aaagggcacatagttacacgtttaaagggtggcgatccatttgtttttggacgtggcgga
gaagaagcagaggaattagcatgtcataatatccactttgaaattatacctggaattcca
gtaacacatcgtgattatagttcttctgtagcatttgtaactgcagtgaataaacctggt
atggataaaggcaaatactggcaacatttggccaatggtcctgaaactttatgtatttat
10   atgggggttaagagactcagtgaaatttgtgagttgttaatacaatatggtcgttcgtca
gaaacaccagtagctctcgtgcatatgggaacgtcaaaacagcaaatgacagtgactggg
acactcgatacaattcaagaacgagcacatcatattcagaatccagcaatgattattgta
ggcgaagtggttaagatgagagagaaaaaattaattggtttgtagaacaggcaactgttcaa
aatgaaacgttaacggaaatgtcatcaacttag

15

Sequence 2174
VVHMGKVYLVGAGPGDPELITLKGLKAIKEADVILYDRLVNKEILNYASPSTKFFYCGKD
PHRHSLPQEETNKMMVTLAKKGHIVTRLKGGDPFVFGRGGEEAEELACHNIHFEIIPGIP
VTHRDYSSSVAFVTAVNKPGMDKGKYWQHLANGPETLCIYMGVKRLSEICELLIQYGRSS
20   ETPVALVHMGTSKQQMTVTGTLDTIQERAHHIQNPAMIIVGEVVKMREKINWFVEQATVQ
NETLTEMSST*

Sequence 2175
Contig_0719_pos_2721_2116,
25   is similar to (with p-value 4.0e-19)
>gp:gp|AJ000974|BSPYREYLO_8 Bacillus subtilis pyrE to yloA g
ene region. NID: g2462954. >gp:gp|Z99112|BSUB0009_34 Bacillu
s subtilis complete genome (section 9 of 21): from 1598421 t
o 1807200. NID: g2633902.
30   atgcccttaatgattgatttaagtaacaagaaagtcgtcattgtaggtggaggtaaagtg
gcaacacgtcgtgctaaaactttattagcttatacaaaacatattcatgttgtaagtcca
acaattaccgatacattacaaaaatatctagaaacgaagcaaatcacttatgaaaagaaa
cacttcgaaccacaagatgttgagaatgctgatgtggtcatcgcggctactaatcaatct
gatgttaacaacgatgtggggggcagctttgtctaagaacgtattatttaatcatgcagga
35   caagcagacctaggtaatgtaacgttccctaatttcttaaaaagagataaattaacaata
agtgtatcaactgatggtgcaagtcctaaaattaggtcaacgaattattaaagatttaaaa
gatacatacaatgaagactattcaatgtatattcagttttttatatgaaagtagacaatat
attaaatcacttaaaattgagccatctgataaacaagcgttactcgagcaaattttgtca
gacaaatatttagatgagaagaagcaacaagatttcatccgatggctaaaatcacaagtc
40   aaatga

Sequence 2176
MPLMIDLSNKKVVIVGGGKVATRRAKTLLAYTKHIHVVSPTITDTLQKYLETKQITYEKK
HFEPQDVENADVVIAATNQSDVNNDVGAALSKNVLFNHAGQADLGNVTFPNFLKRDKLTI
45   SVSTDGASPKLGQRIIKDLKDTYNEDYSMYIQFLYESRQYIKSLKIEPSDKQALLEQILS
DKYLDEKKQQDFIRWLKSQVK*

Sequence 2177
Contig_0719_pos_2029_1196,
50   putative peptide of unknown function
atgggatttggcgcttcatcgtcatcaatattattaacttacggtatagcaccggcagta
gtgtcagcaaccgttcattttctgaaattgcaacaacagctgcatctgggacatcacat
tggagatttgataatgttcataaaccaacaatgttgaagttagctataacctgggtcaata
agcgcctttatcggtgcaggtgttttgacatttattcatggtgattatattaaaccattc
55   attgctttattcttgttaagtatgggattttatattttgtatcaatttctatttaaacgt
gcacatgaacatcatcatcatgtgggaaatttgagtagttttaaagtaattccacaaggt
tttgtggcaggattttttagacgcaatcggtggtggtggttggggaccggttaatacgccg
ctcctgctttcaagtaaaaaaattcaaccacgatatgcgattggaacagtctcagcaagt
gaattttttgttacgtcatctgccgcctttaagtttcattatcttttttaggagtcactcaa

attaattggtttgctgtaattgctttaagtctcggtggaatggtagcagcacctatttca
gcgtatttagttaaagtgttacccattaacattcttgcaatttgtgtcggtggtttaatt
atttttacaaatagtaatgcattattaagctattttgtaaaagataacactatttcaaat
acagttcgattcattattattcttgcaattattattttgcttgtttttcaagtcgttcga

5   aacaagaaattgtctttttcttataagaaaagccgagtaaacaaatataattaa

Sequence 2178
MGFGASSSSILLTYGIAPAVVSATVHFSEIATTAASGTSHWRFDNVHKPTMLKLAIPGSI
SAFIGAGVLTFIHGDYIKPFIALFLLSMGFYILYQFLFKRAHEHHHHVGNLSSFKVIPQG

10   FVAGFLDAIGGGGWGPVNTPLLLSSKKIQPRYAIGTVSASEFFVTSSAALSFIIFLGVTQ
INWFAVIALSLGGMVAAPISAYLVKVLPINILAICVGGLIIFTNSNALLSYFVKDNTISN
TVRFIIILAIIILLVFQVVRNKKLSFSYKKSRVNKYN*

Sequence 2179
15   Contig_0719_pos_0_1172,
is similar to (with p-value 0.0e+00)
>gp:gp|AJ000974|BSPYREYLO_4 Bacillus subtilis pyrE to yloA g
ene region. NID: g2462954. >gp:gp|Z99112|BSUB0009_30 Bacillu
s subtilis complete genome (section 9 of 21): from 1598421 t

20   o 1807200. NID: g2633902.
atgtctaacaatgaaacaataaccaattatacaattaaacctcatggaggagaactcatc
aatcgtgttgttgaaggaaacgaacgtgaacgtttgattgaggaagcattaaattttaaa
ccgattactttaaatccttggggaatatcggatctagagctcataggtattggcggattt
agtcccccttacaggatttatgaacaaggaagactcaaggttatagaggaaacacat

25   ttaagcaatggcttagtttggagtattcctatcacttacctgtaacagaatccgaagca
gataaacttgaaataggtgatgatattgctttatatggtgaagatggtcagttatatgga
acgcttaaattagaagaaaagtacacatatgataaagaaaaagaagcgcgtttggtgtac
ggaactactgaagaagctcatcctggagttaaaaaggtttatgaaaaaggtaatatatat
ttaggtggtcctattaaactattaaatcgtccaaaacatgacgcgttttcaaattatcat

30   ctggatccttcagagacgagacaattatttcatgatttaggttggaaaactgtcgtaggt
tttcaaacgagaaatccagtgcatcgagcacatgaatatattcaaaaatcagcactagaa
attgttgatggtctacttttaaatccactagttggtgaaacaaagtcagacgatattcca
gcggatgtacgtatggaaagttatgaagtgatattaaaaaaactattatcctgaagataga
gcacgtctagtcatttatcctgctgcaatgcgctatgccggaccacgtgaagcgatactt

35   catgcaactgtccgtaaaaattatggttgtacacatttattgtgggaagagatcacgct
ggggtaggcgattattatggtacttatgaagcacaagagctgattactcaatttgaagat
gagttaggtattcaaattttaaaatttgaacatgccttttattgcgaagcttgtggtaat
atggcaactgctaaaacatgtccgcatgacgcttctcaacatttacatttaagtggtact
aaagtaagagaaaaactgcgtaatggcgaatcattgccaactaaattttcaagaccagaa

40   gttgccgaagttctaattaaaggtttgcgagT

Sequence 2180
MSNNETITNYTIKPHGGELINRVVEGNERERLIEEALNFKPITLNPWGISDLELIGIGGF
SPLTGFMNKEDYTKVIEETHLSNGLVWSIPITLPVTESEADKLEIGDDIALYGEDGQLYG

45   TLKLEEKYTYDKEKEARLVYGTTEEAHPGVKKVYEKGNIYLGGPIKLLNRPKHDAFSNYH
LDPSETRQLFHDLGWKTVVGFQTRNPVHRAHEYIQKSALEIVDGLLLNPLVGETKSDDIP
ADVRMESYEVILKNYYPEDRARLVIYPAAMRYAGPREAILHATVRKNYGCTHFIVGRDHA
GVGDYYGTYEAQELITQFEDELGIQILKFEHAFYCEACGNMATAKTCPHDASQHLHLSGT
KVREKLRNGESLPTKFSRPEVAEVLIKGLRV

50
Sequence 2181
Contig_0720_pos_409_906,
is similar to (with p-value 3.0e-25)
>sp:sp|P35164|RESE_BACSU SENSOR PROTEIN RESE (EC 2.7.3.-). >

55   pir:pir|S45560|S45560 hypothetical protein X18 - Bacillus su
btilis >gp:gp|L09228|BACDIA_27 Bacillus subtilis spoVA to se
rA region. NID: g410114. >gp:gp|Z99116|BSUB0013_23 Bacillus
subtilis complete genome (section 13 of 21): from 2395261 to
2613730. NID: g2634723.

```
atggatgctgaaggattatcagttgagaaggaattacaacctattcaacaccttcttgat
aaaatggagtctaaatatcgcatgcaaagtgaagaattaggtttaacaatgacgtttgat
tctaataatgacgaacaattatggaactatgatatggatagaatggaccaagtgttaact
aatttaattgataacgcaacaagatatacacaagctggtgattctataaagatttctatt
gatgaagattcagatttcaatattttaacaataactgatacaggcactggtatagcaccg
gaacatctgaaacaagtatttgaccgttttttataaagtggacgctgctcgaaaaagaggt
aagcaaggcaccggattaggactttcatttgtaaaatgattattgaagaacacgggggga
cgtattgatgttgagagcgaattaggcaaaggtacttcatttattattagactacctaaa
tcaaacaaattagttag
```

Sequence 2182
MDAEGLSVEKELQPIQHLLDKMESKYRMQSEELGLTMTFDSNNDEQLWNYDMDRMDQVLT
NLIDNATRYTQAGDSIKISIDEDSDFNILTITDTGTGIAPEHLKQVFDRFYKVDAARKRG
KQGTGLGLFICKMIIEEHGGRIDVESELGKGTSFIIRLPKSKQIS*

Sequence 2183
Contig_0720_pos_1420_2037,
is similar to (with p-value 7.0e-39)
>sp:sp|P50726|YPAA_BACSU HYPOTHETICAL 20.5 KD PROTEIN IN SER
A-FER INTERGENIC REGION. >gp:gp|Z99116|BSUB0013_17 Bacillus
subtilis complete genome (section 13 of 21): from 2395261 to
 2613730. NID: g2634723. >gp:gp|L47648|BACSERA_2 Bacillus su
btilis phosphoglycerate dehydrogenase (serA), ypaA, ferredox
in (fer), ypbB, recS, ypbD, ypbE, ypbF, ypbG, ypbH, glutamat
e dehydrogenase (ypcA), ypdA, ypdB, ypdC, spore cortex lytic
 enzyme (sleB), ypeB, ypfA, ypfB, cytidine monophosphate kin
ase (cmk), ypfD, ypgA, yphA, yphB, yphC, NAD+ dependent glyc
erol-3-phosphate dehydrogenase (glyc), yphE and yphF genes,
complete cds. NID: g1146195. >gp:gp|L47648|BACSERA_2 Bacillu
s subtilis phosphoglycerate dehydrogenase (serA), ypaA, ferr
edoxin (fer), ypbB, recS, ypbD, ypbE, ypbF, ypbG, ypbH, glut
amate dehydrogenase (ypcA), ypdA, ypdB, ypdC, spore cortex l
ytic enzyme (sleB), ypeB, ypfA, ypfB, cytidine monophosphate
 kinase (cmk), ypfD, ypgA, yphA, yphB, yphC, NAD+ dependent
glycerol-3-phosphate dehydrogenase (glyc), yphE and yphF gen
es, complete cds. NID: g1146195.
```
atgggagaagatggaggttttttgttgttcaaatttcctcctttctaagtaagatgaat
ggaaggagaaaattatatatgcaacaaaacaaacgtttgattacaattagtatgttaagt
gcggtagcgtttgtgttaactttcatcaagtttccattgccatttataccaccgtatcta
actctcgattttagtgatgtaccgacgttattagcaacattcctcttaagtcctattgct
gggattatcgttgcactcatcaaaaatattttaaattttctattcaatataggggatcct
gttggaccagtagctaactttttagcaggcgtcagcttttttgctatcatcatactatgtt
tatagaaaaagaaaaaataatcgttctttaatttatggattaattacaggtacaatcgtt
atgactattgttttgagcatcttaaattattttgtgttacttccattatatggaatgata
tttaatttaggtgatgtgcttaataacgtaaaaattgttattgtgtctggagtcatacct
tttaatttaattaaaggcataatcatttccattatatttgtgctgttatttagaagatta
agacatatcatcaaataa
```

Sequence 2184
MGEDGGFLLFKFPPFLSKMNGRRKLYMQQNKRLITISMLSAVAFVLTFIKFPLPFIPPYL
TLDFSDVPTLLATFLLSPIAGIIVALIKNILNFLFNIGDPVGPVANFLAGVSFLLSSYYV
YRKRKNNRSLIYGLITGTIVMTIVLSILNYFVLLPLYGMIFNLGDVLNNVKIVIVSGVIP
FNLIKGIIISIIFVLLFRRLRHIIK*

Sequence 2185
Contig_0720_pos_2823_3761,
putative peptide of unknown function
```
atgagtcatgcatttaactacaaaaccaataaaagtatctataatattttaacaggcaag
aagtcacaccaaacgtttttcgatgcgtcaagccaacaacttttgtcattatatcatagt
```

```
ttacctaacttaaagtattcaacttttgagcaatttatactccaaaaggatgatttaaa
aaatcaattcaagtaaaaatacatccacaatacacttatgatagtctcactcaaaccttt
agttgcatacaattacttatccaaacgttatctcatacacgcaaggagtcaaatacattt
attccaatcgttcaaaataccctatatccaacaaagggttaaacaactttatcatcaggtc
```

5
```
attgaatcaaatcaagtatcaaatactatagacgaaatatatttattatttgagaattta
aataataaatataaccatacatttcttcattattatttacaaggatatgaggaatccatg
tatactagacaacaaataagtttaattgagagtataccacaatcagaattatttgaacga
gaaatgaatgaactgattgacatattgaatcaattaaaagattcaacgaaatatccaata
ctttctcaagctatcattctttcaccattactaacaaatacatacttaagctatcaaaag
```
10
```
ttaaaatctggtctcaatttaaaagaaattgctcaattacaaaatgtaaaacttaacaca
attgaagatcatattctagaaatgtatattaaaggttatttgatagactatacattattt
ataaataaaaaagatattctcgaatttataaactactatcaaaaacatcgcggtgaacga
ttaaaattttataaagaacattttactgattggacttactttcaaattaagttagttata
gtaggaatagaaagaggtgatttaattgctgaaagataa
```
15

Sequence 2186
MSHAFNYKTNKSIYNILTGKKSHQTFFDASSQQLLSLYHSLPNLKYSTFEQFILQKDDFK
KSIQVKIHPQYTYDSLTQTFSCIQLLIQTLSHTRKESNTFIPIVQNTYIQQRVKQLYHQV
IESNQVSNTIDEIYLLFENLNNKYNHTFLHYYLQGYEESMYTRQQISLIESIPQSELFER
20 EMNELIDILNQLKDSTKYPILSQAIILSPLLTNTYLSYQKLKSGLNLKEIAQLQNVKLNT
IEDHILEMYIKGYLIDYTLFINKKDILEFINYYQKHRGERLKFYKEHFTDWTYFQIKLVI
VGIERGDLIAER*

Sequence 2187
25 Contig_0720_pos_5757_5161,
putative peptide of unknown function
atgtcaatggataaagctacaaaatatgcacttcaattaagggtgattgctcaagaaagc
tatcctagtgtcgatagacattcctttctagtcgaaaaattacgcttacttattcaacaa
ttaaaacaatctattcatcatctcaaacaattagatgatgccatgattcaattagcacaa
30 caactcgattattttgaaaatattcattcgatacctggtattggtaagctaagcacagct
atgattattgggggagattggtgatattaagcgatttaaatcaaataaacaactcaatgct
tttgttggcattgatatcaaacgatatcaatcaggtcatacacactgtagagataccatc
aacaagcgtggtaataaaaaagcgagaaaacttttattttgggtgattatgaatataata
agagggcagcatcattatgacaatcatgtcgtcgattattactacaaactaagaaagcag
35 cctaatgagaaacctcataagactgccatcattgcttgtataaatcgattattaaaaaca
attcattatcttgtaatgaatcataaattgtacgattatcaaatgtcaccacattag

Sequence 2188
MSMDKATKYALQLRVIAQESYPSVDRHSFLVEKLRLLIQQLKQSIHHLKQLDDAMIQLAQ
40 QLDYFENIHSIPGIGKLSTAMIIGEIGDIKRFKSNKQLNAFVGIDIKRYQSGHTHCRDTI
NKRGNKKARKLLFWVIMNIIRGQHHYDNHVVDYYYKLRKQPNEKPHKTAIIACINRLLKT
IHYLVMNHKLYDYQMSPH*

Sequence 2189
45 Contig_0721_pos_858_1772,
is similar to (with p-value 0.0e+00)
>sp:sp|P50307|METK_STAAU S-ADENOSYLMETHIONINE SYNTHETASE (EC
  2.5.1.6) (METHIONINE ADENOSYLTRANSFERASE) (ADOMET SYNTHETAS
  E). >gp:gp|U36379|SAU36379_1 Staphylococcus aureus S-adenosy
50 lmethionine synthetase gene, complete cds. NID: g1020316.
atggttaaaaataaacagtctcctgatattgcacagggtgtagacaaagctcttgagtat
cgaaatgatatttctgaagaagaaattgaagcaacaggtgcaggtgaccaaggattaatg
tttggatatgcaactgatgaaactgatacgtatatgcctctacctatattcctgtcacat
caacttgctaaacgattggctgatgtacgaaaagatgaaattttagattatcttcgtcca
55 gatggaaaagtacaggtgactgttgaatatggtaaagatgacaaacctagacgtattgat
accattgtagtttctacacaacatgctgaagatgtagagttagcacaaattgaaaaggac
attaaaacgcatgttatttacccaactgtagataaagctttattagatgatgaaactaaa
ttttacattaacccaactggacgtttcgttattggaggacctcaaggagatgctggttta
actggacgtaaaattatagttgatacgtatggtggttatgcccgtcatggtggaggttgt
```

```
     tttagtggtaaagatcctactaaagtagatcgttcagcagcttatgcagcaagatatgta
     gctaaaaatattgttgcagctggtttagctaaacaatgtgaagtacaacttgcatatgca
     attggtgtagcagaacccgtttccatttcaattaatacgtttgatactggaaaggtttca
     gaagcacgtttagttgaagctgtaagaaagcattttgatttaagaccagcaggtatcatt
  5  aaaatgttagacttaaaacaaccgatatatagacaaacagcagcgtatggtcattttgga
     cgtacagacgtattgttaccatgggaaaaattagataaagtcaatgtttttaaaagatgct
     gttgaaattcaatga


     Sequence 2190
 10  MVKNKQSPDIAQGVDKALEYRNDISEEEIEATGAGDQGLMFGYATDETDTYMPLPIFLSH
     QLAKRLADVRKDEILDYLRPDGKVQVTVEYGKDDKPRRIDTIVVSTQHAEDVELAQIEKD
     IKTHVIYPTVDKALLDDETKFYINPTGRFVIGGPQGDAGLTGRKIIVDTYGGYARHGGGC
     FSGKDPTKVDRSAAYAARYVAKNIVAAGLAKQCEVQLAYAIGVAEPVSISINTFDTGKVS
     EARLVEAVRKHFDLRPAGIIKMLDLKQPIYRQTAAYGHFGRTDVLLPWEKLDKVNVLKDA
 15  VEIQ*


     Sequence 2191
     Contig_0721_pos_1923_2783,
     putative peptide of unknown function
 20  gtggctgttttattctttatactcttcttagtagcgaaccatagtaagaagaaagtgaag
     aatcaaacagaagcacattataaagaaaaagaacaacatctaaaagaatctcatgaagaa
     gctttagaaaaagagagagttgagaataaaaaagttgttacaaaacaaaaagaagatttt
     gacgtgacagttagtaacaaaaatcgtgaaattgatgcgttgaaactattctcaaaaaat
     catagtgaatatgttacagatatgcgattaattggtattcgtgagagactagtaaatgaa
 25  aagagaatacgccctgaagatatgcatattatggcaaatattttcttgcctagtaatgag
     ttaactaatattgaacgtgtgagtcatcttgtacttacacgaactggcttatatataatt
     gattctcaattattaaaaggccatgtttataatggtattagtggtgcgcaatttaaagaa
     ttacctacgatgtcacaagttttttgacacgctcgacttagattcatcacagccacagaca
     ttggtcttagatcaaaatgaagatcaacattcattatctttttgttaattattcagataag
 30  attaaacatattgaaaaattagcaggagatttacaaaatgaattgaataccaaatatacg
     cctacatcaatactgtattttaatcctaaaaaggataatgatgttacaatttctcattat
     acgcagtcatcaaatgttaaagtttttagttggtcctgaacaattagatgaattctccaac
     aagtttgttttccatggacgcatacagtacaatgtggatgatttacaagatatcatggat
     aaaatcgagtcattcaattaa
 35

     Sequence 2192
     VAVLFFILFLVANHSKKKVKNQTEAHYKEKEQHLKESHEEALEKERVENKKVVTKQKEDF
     DVTVSNKNREIDALKLFSKNHSEYVTDMRLIGIRERLVNEKRIRPEDMHIMANIFLPSNE
     LTNIERVSHLVLTRTGLYIIDSQLLKGHVYNGISGAQFKELPTMSQVFDTLDLDSSQPQT
 40  LVLDQNEDQHSLSFVNYSDKIKHIEKLAGDLQNELNTKYTPTSILYFNPKKDNDVTISHY
     TQSSNVKVLVGPEQLDEFFNKFVFHGRIQYNVDDLQDIMDKIESFN*


     Sequence 2193
     Contig_0721_pos_3101_3799,
 45  is similar to (with p-value 2.0e-56)
     >gp:gp|AL034443|SC4B5_1 Streptomyces coelicolor cosmid 4B5.
     NID: g4007668.
     atgatttatggtaacgaagaaacggtaggtcaaggtattaaagaaggattagaatcgact
     ggtttaagtcgtgaagatttatttataacttctaaattatggctaactgattttggacgt
 50  caaaatgtggaagatgcctatcgacaatctgttgcaaaattaggactggattatttagat
     ttatatctgatgcattggccaggtacaaatgaagcggtaatgattgatacttggacaggt
     atggaagacttgtataaacaaaatcaggtgaaaaatattggtgtgagcaattttatccca
     gaacattttgaagcattgcttgcccaagtttctattaaaccggtcattaaccaagtagaa
     tttcatccttatttaacacaaaatgaattacgaaagtatttagaagctcaaaatattatc
 55  atggaatcatggtctccattaatgaattcacaaattctccatgatgaagtcataaatgaa
     gtagctaatgaagtaggaaaaactccagcccaagttgtaataagatggaatattcagcac
     gatgttgttgttatacctaaatccgtaacaccacatcgcatagaagaaaatttagacgta
     tggaattttgaattaagcgacaaccaaatggaagaatcgatcaattaaatcaagataaa
     agaattggacctaacctcttgaatttaacggtaagtaa
```

Sequence 2194
MIYGNEETVGQGIKEGLESTGLSREDLFITSKLWLTDFGRQNVEDAYRQSVAKLGINDYLD
LYLMHWPGTNEAVMIDTWRGMEDLYKQNQVKNIGVSNFTPEHFEALLAQVSIKPVINQVE
5   FHPYLTQNELRKYLEAQNIIMESWSPLMNSQILHDEVINEVANEVGKTPAQVVIRWNIQH
DVVVIPKSVTPHRIEENLDVWNFELSDNQMERIDQLNQDKRIGPNPLEFNGK*

Sequence 2195
Contig_0721_pos_5489_4476,
10   putative peptide of unknown function
atggaacgattttgttgtgtaaatcaaattaactatattcaaatgaatccgttagaagcc
aaatttaaaacgagcgctctaagatcatggaaaactgatcaggcagatgctcataagctt
gcttgtttaggaccgacgcttaaacaaacagacagcttacctatacatgagttaatattc
tttgaattaagagaacgcgtccgttttcatctagaaatcgagaatgaacaaaatcgactt
15   aaatttcagatccttgaattactccatcaaacattccctggtttagaaagattgtttagt
agtcgatattcaatcattgcactcaacatcgcagaaatctttactcatccagacatggtt
cttgatatcgacaaggatgtactgattacacatatattcaattctacagataagggaatg
tcaatggataaagctacaaaatatgcacttcaattaagagtgattactcaagaaagctat
cctaatgtcgatagacattcctttctagtcgaaaaattacgcttacttattcaacaatta
20   aaacaatctattcatcatctcaaacaattagatgatgccatgattcaattagcacaacaa
ctcgattattttgaaaatattcattcgatacctggtattggtaaactaagcacagctatg
attattggggagattggtgatattaagcgatttaaatcaaataaacaactcaatgctttt
gttggcattgatatcaaacgatatcaatcaggtcatacacactgtagagataccatcaac
aagcgtggtaataaaaaagcgagaaaacttttattttgggtgattatgaatataataaga
25   gggcagcatcattatgacaatcatgtcgtcgattattactacaaactaagaaagcagcct
aatgagaaacctcataagactgccatcattgcttgtataaatcgattattaaaaacaatt
cattatcttgtaatgaatcataaattgtacgattatcaaatgtcaccacattag

Sequence 2196
30   MERFCCVNQINYIQMNPLEAKFKTSALRSWKTDQADAHKLACLGPTLKQTDSLPIHELIF
FELRERVRFHLEIENEQNRLKFQILELLHQTFPGLERLFSSRYSIIALNIAEIFTHPDMV
LDIDKDVLITHIFNSTDKGMSMDKATKYALQLRVITQESYPNVDRHSFLVEKLRLLIQQL
KQSIHHLKQLDDAMIQLAQQLDYFENIHSIPGIGKLSTAMIIGEIGDIKRFKSNKQLNAF
VGIDIKRYQSGHTHCRDTINKRGNKKARKLLFWVIMNIIRGQHHYDNHVVDYYYKLRKQP
35   NEKPHKTAIIACINRLLKTIHYLVMNHKLYDYQMSPH*

Sequence 2197
Contig_0724_pos_3443_3748,
putative peptide of unknown function
40   atgctcctttctgctatactccttataaaaggaggtgaaaatatgaaaagtttattatt
gcgtatgatttaaataaccaaaaggattatccaaaattaatagagcgtattgaggattat
cctaatgttgctaaaatcaataaatcagtttggtttattaattcaactaatgatgctaaa
actatttagaaacgaattaaaaatgtttattgatagcgatgatagtttgttcgttggtaag
ctgactggtgaagccgcatggtctaatgtaatttgcagttcacaacatttaaaagattat
45   ctttag

Sequence 2198
MLLSAILLIKGGENMKSFIIAYDLNNQKDYPKLIERIEDYPNVAKINKSVWFINSTNDAK
TIRNELKMFIDSDDSLFVGKLTGEAAWSNVICSSQHLKDYL*

50
Sequence 2199
Contig_0724_pos_4321_4635,
putative peptide of unknown function
atgtgcttttcaaaaagaatgaaacaatcaagagaaaaacaaggtatgactttagctgaa
55   ctaggaagaaaaatcggtaaaactgaagctactgtacaacgttatgaaagcgggaatatt
aaaaatcttaaaaatgatactattgaaagtatagctactgcattaaatgttaaccctgct
ttcttgatgggttggatagaagaagttgaggaacaaccacaacatcgtgcagcgcatctt
gatggtgatttaactgacgaagaatggcaagaaattcttgattacgctgaatacataaga
agtaaaagaaaataa

Sequence 2200
MCFSKRMKQSREKQGMTLAELGRKIGKTEATVQRYESGNIKNLKNDTIESIATALNVNPA
FLMGWIEEVEEQPQHRAAHLDGDLTDEEWQEILDYAEYIRSKRK*

Sequence 2201
Contig_0724_pos_4639_0,
putative peptide of unknown function
gtgttttatgtggggaaatatgaagatatgttaattgaacatgactatattgaagtcatt
gaatgtgataacttacctaaaaggttatctggtttgtggcttggagatatgattttaatt
aatcgtaacttgcctattacttccaaacttgaaacacttgcagaggaactcgctcataac
gaacttacatatggaaatatagttgatcaaagtagttttaatcatagaaaatttgaaggt
tatgcacgtaggttagcctatgaaaagttaatccctcttaaagatattgtaaaagcattt
ttccaaggcattcatgacttgtatgaacttgctaattttttga

Sequence 2202
VFYVGKYEDMLIEHDYIEVIECDNLPKRLSGLWLGDMILINRNLPITSKLETLAEELAHN
ELTYGNIVDQSSFNHRKFEGYARRLAYEKLIPLKDIVKAFLQGIHDLYELANFFX

Sequence 2203
Contig_0724_pos_3437_2028,
is similar to (with p-value 0.0e+00)
>gp:gp|M57689|BACSPO0K_5 Bacillus subtilis spo0K operon. NID
: g143602.
atgcaagatttacaagtatttaattttgaagatttaccagtaagaaaaatagaagtagat
ggagaaccatatttttttaggtaaagacgtggcagaaatattaggttacacaagatctgat
aatgcaattagaaatcatgttgatgatgaagataagctgacgcaccaagttagtgcatca
ggtcaaaaacgaaacatggtaatcatcaacgaatctggtttatacagcttaatctttgac
gctgctaaacaaagtaaaaacgaaagtattagaaagaaagctaaacgttttaaacgtccc
aaaaatgcgaaaataaaggaggtaaaagtaatgacagaaacggtattagaagtaaatgat
ttgcacgtttcctttgatattgctgcaggagaagtgcaagctgtcagaggcgtggatttt
catttaaataaaggggaaacgttagccattgttggagaatctggatctggaaaatctgta
acaactaaggcaattacaaaacttttttcaaaaggatacaggaagaataaaaaaagggagaa
atcttattttttaggtgaggacttagctcagaaagtgaaaaagaactgatacagctaaga
ggtcgagatatttcaatgatatttcaggatcctatgacttctttaaatccaacaatgcaa
atcggaaagcaagtcatggaacctttgattaaacataagaaattaagtaaagcaaaggcc
aagcaaagagcattggaaattttgaatttagttgggttacctcgtgctgaaaaacgattt
aaagcttatccacatcaatttttcaggaggacaacgtcagagaatagttattgcaatagca
ttggcatgtgagcctaaaatattaattgctgatgagcctacaactgctttagatgtgaca
atgcaggctcaaatttttagatcttatgaaagaactacaaaataagattgaaacttcaatt
atctttattacgcatgatttaggcgttgtagcaaatattgcggacaaagtagccgtaatg
tatgggggacagatggttgaaacaggggatgtgaatgaaatattttatgatcctaaacat
ccctatacctggggattgctttcttcaatgcctgatttaacaaccagtaatgacacggac
ttaattgcaattccaggtacaccaccagatttacttcatccaccaattggtgatgctttc
gcacgtagaagtcgatatgctttagatattgatttttaaagaagaaccaccttggttcaaa
atttcacccacacattttgttaaatcttggttattagatgcaagagctccaaaagttacg
ccaccttcaatggttcaaaaacgattaagaacaatgccaagtaattatgaacaaccacat
agagtagagagggtggctttttaatgagtaa

Sequence 2204
MQDLQVFNFEDLPVRKIEVDGEPYFLGKDVAEILGYTRSDNAIRNHVDDEDKLTHQVSAS
GQKRNMVIINESGLYSLIFDAAKQSKNESIRKKAKRFKRPKNAKIKEVKVMTETVLEVND
LHVSFDIAAGEVQAVRGVDFHLNKGETLAIVGESGSGKSVTTKAITKLFQKDTGRIKKGE
ILIYGEDIAQKSEKELIQLRGRDISMIFQDPMTSLNPTMQIGKQVMEPLIKHKKLRKAKA
KQRALEILNLVGLPRAEKRFKAYPHQFSGGQRQRIVIAIALACEPKILIADEPTTALDVT
MQAQILDLMKELQNKIETSIIFITHDLGVVANIADKVAVMYGGQMVETGDVNEIFYDPKH
PYTWGLLSSMPDLTTSNDTDLIAIPGTPPDLLHPPIGDAFARRSRYALDIDFKEEPPWFK
ISPTHFVKSWLLDARAPKVTPPSMVQKRLRTMPSNYEQPHRVERVAFNE*

547

Sequence 2205
Contig_0724_pos_1756_1097,
is similar to (with p-value 3.0e-81)
>pir:pir|E38447|E38447 sporulation initiation protein spo0KE
5   - Bacillus subtilis >gp:gp|M57689|BACSPO0K_6 Bacillus subti
lis spo0K operon. NID: g143602.
atgatatttcaagatccttatgcatctctaaatcctcggttaaaggtaatggatatcgta
gctgaaggaatagatattcacaaacttgctagtagtcagcgtgatcgaaagaaacgtgta
tacgacctttttagaaacagttggtttaggtaaagaacacgcgaatcgttacccacatgag
10   ttttcaggcggacaaagacaacgtattggtatcgcacgtgcattagctgtagagccagaa
tttattattgcagatgaaccgatatcagcattagatgtatcgattcaagctcaagtcgtt
aatcttttattaaagctacaacgtgaacgtgatattactttattgtttattgctcatgat
ttatcaatggtgaaatatatttccgatagaattgcagtgatgcacttcggtaaaattgta
gaaattggaccggctgatgatatttataattatccattacatgattatactaagtcatta
15   ttaagtgccattccacagcctgatcctgatgttgagagaaatcgtcaacgtgtttttatat
catgaagatgcaacgcttaatgaagaacgtcaattaaatgaaattagaccacaacattat
gtattttctactcaaaacgaagcagttaaattgaaacaaaagtatggtttgtctgtttaa

20   Sequence 2206
MIFQDPYASLNPRLKVMDIVAEGIDIHKLASSQRDRKKRVYDLLETVGLGKEHANRYPHE
FSGGQRQRIGIARALAVEPEFIIADEPISALDVSIQAQVVNLLLKLQRERDITLLFIAHD
LSMVKYISDRIAVMHFGKIVEIGPADDIYNYPLHDYTKSLLSAIPQPDPDVERNRQRVLY
HEDATLNEERQLNEIRPQHYVFSTQNEAVKLKQKYGLSV*
25

Sequence 2207
Contig_0724_pos_1070_234,
is similar to (with p-value 2.0e-51)
>pir:pir|A53310|A53310 pheromone cAD1 binding protein precur
30   sor - Enterococcus faecalis plasmid pAD1 >gp:gp|L19532|AD1TR
AC_2 Plasmid pAD1 (from Enterococcus faecalis strain: DS16)
hemolysin bacteriocin (traC) gene, complete cds, traA and tr
aB genes, 3' end. NID: g388267.
atgaaaggtttttaaagtcttaattattttattaagtgtatgcataattttatctgcttgt
35   agtaataagcagagtttatattcagaccaggggcaagttttttaggaaggtaatcacacaa
gatatgactacactagatacagctttaattacagatgctgtttctggtgatatagcagct
caagcttttgaaggattatatactttaaataaagaagacaaagctgaaccagctattgct
aaatcttttccaaagaaaagtaatggtggcaaaacacttacgattaatttaagaaaaaat
gcaaaatggtccaatggagattcggtaactgcatatgacttcgtatatgcgtggagaaag
40   gtagttaatcctaagacggcttctgagtttgcatacataatgagcgatataaaaaatgca
gatgaagttaatgcaggtaaaaaatcagtcaaggatttgggtatcaaggctataggtaaa
tataaattacaagtagatttagaaagacctgtaccttatattaatgaactattagcactt
aatacatttaatcctcaaaatgagaaagttgctaaaaagtttggagaacaatatggtaca
actgctgaaaaagcagtgtacaatggaccatttgaagtaacaaattggaaagtggaagat
45   aaaattcaattagttaaaaatgaacaatattgggataagaagaatgtaaaattagataaa
gtgaactataaagtattaaaagatcaacaagcaggtgcatcgttatatgatactggctcg
gtcgatgatactatgttaagtatactgcacaagcatctagtccatcagaaggtttag

Sequence 2208
50   MKGFKVLIILLSVCIILSACSNKQSLYSDQGQVFRKVITQDMTTLDTALITDAVSGDIAA
QAFEGLYTLNKEDKAEPAIAKSFPKKSNGGKTLTINLRKNAKWSNGDSVTAYDFVYAWRK
VVNPKTASEFAYIMSDIKNADEVNAGKKSVKDLGIKAIGKYKLQVDLERPVPYINELLAL
NTFNPQNEKVAKKFGEQYGTTAEKAVYNGPFEVTNWKVEDKIQLVKNEQYWDKKNVKLDK
VNYKVLKDQQAGASLYDTGSVDDTMLSILHKHLVHQKV*
55

Sequence 2209
Contig_0727_pos_3292_0,
putative peptide of unknown function
atgtcttttcttaggaaacacaccgaaattatatttagttatatcatcggtatcgtttca

```
cttttttacaggtctcattattttttattaacttacctttaatcaaacaatttaaaggtgac
aaaaaggttgatacgcatgtgcataacgtatgggaattcctaaatgcctttttttgccgag
attataaaagtgatgagtaaatttattggtggctttccaattacaagtgccatagtaatc
atcgtatttggtattctagtgatgctgttaggtcacactttatttagaactattaaatac
```
5
```
gattatgacatttcaattttctttttagttattggcattatgtactttatcattacatta
ttgctaatgacacaagtgtatggcttttttgctatcgtctttattattccttttacagtt
catattggttacatagtttataaagatgagttgaaccaagacaatcgaaagaaccattat
atgtggattattgtaacttatggaatgagttatcttattacccaaatttcgctatatgga
cgtattgacgcaaatgaaattgaatcaattgatattttaagtgtaaatacattcttcatt
```
10
```
attatgtggttattaggtcagatggctata
```

Sequence 2210
MSFLRKHTEIIFSYIIGIVSLFTGLIIFINLPLIKQFKGDKKVDTHVHNVWEFLNAFFAE
IIKVMSKFIGGFPITSAIVIIVFGILVMLLGHTLFRTIKYDYDISIFFLVIGIMYFIITL
15
LLMTQVYGFFAIVFIIPFTVHIGYIVYKDELNQDNRKNHYMWIIVTYGMSYLITQISLYG
RIDANEIESIDILSVNTFFIIMWLLGQMAI

Sequence 2211
Contig_0727_pos_0_369,
20
putative peptide of unknown function
```
atgtttccaccatatttttatcacgaacagggagtgtaagagcatgacgcgacagagaatc
gccattgatatggatgaagtgcttgctgatacattgggtgctgttgttaaagcggtcaat
gaacgagcggatttaaatatcaaaatggaatcattaaacggtaaaaaattaaaacatatg
atacccgagcatgaggggttagtcatggatattttaaaagaacctggattctttagaaat
```
25
```
ttagatgtaatgccgcatgctcaagaagttgtaaaacaactcaatgagcattacgacata
tacatagccacagcagcgatggatgttccaacctcttttcatgacaaatatgaatggtta
ttCAAATGA
```

Sequence 2212
30
MFPPYFITNRECKSMTRQRIAIDMDEVLADTLGAVVKAVNERADLNIKMESLNGKKLKHM
IPEHEGLVMDILKEPGFFRNLDVMPHAQEVVKQLNEHYDIYIATAAMDVPTSFHDKYEWL
FK*

Sequence 2213
35
Contig_0728_pos_6214_6576,
is similar to (with p-value 4.0e-29)
>gp:gp|AJ000339|LDGAPPGK_3 Lactobacillus delbrueckii ygaP, g
ap, pgk, tpi, and ycsE genes. NID: g2624189.
```
atgactttctattgtttttattcccacatcaaagtatttaaaaacttatttatcattaata
gcttttataccaggtaattcttttaccttctaagtattctaatgatgctcctccaccagta
```
40
```
gagatgtgtgtaaagtcatcttcgaaacctaatgaaattgctgctgcggcagagtcacca
ccaccaataatagtagtagcgtcttccaatttagcaatagactcacatacaccgattgta
cctttagcaaaattactaaattcgaatacacccataggtccattccatactacagtatgt
gcaccttgtaattctttattaaataattctactgttttaggtccaatatccattgcttct
```
45
```
tga
```

Sequence 2214
MTFYCFIPTSKYLKTYLSLIAFIPGNSLPSKYSNDAPPPVEMCVKSSSKPNEIAAAAESP
PPIIVVASSNLAIDSHTPIVPLAKLLNSNTPIGPFHTTVCAPCNSLLNNSTVLGPISIAS
50
*

Sequence 2215
Contig_0728_pos_6923_6258,
is similar to (with p-value 2.0e-69)
55
>gp:gp|AJ000339|LDGAPPGK_3 Lactobacillus delbrueckii ygaP, g
ap, pgk, tpi, and ycsE genes. NID: g2624189.
```
atggaaaagaaattaaatttattggtggcgtagtgaatgatccacaaaaaccagtagtt
gctattttaggtggcgctaaagtttcagataaaattaatgttatcaaaaatttagttaat
atcgcagataaaatcttaatcggtggcggtatggcttatacatttattaaagcgcaaggt
```

```
aaggaaataggtctttcattattggaagaagacaaaattgattttgctaaagacttgtta
gagaataatggcgatcaaatagtattacctgtagattgtaaaatcgctaaagaattttct
aatgatgcaaaaatcactgaagtatctatcaatgaaatcccttcagatcaagaagcaatg
gatattggacctaaaacagtagaattatttaataaagaattacaaggtgcacatactgta
gtatggaatggacctatgggtgtattcgaatttagtaattttgctaaaggtacaatcggt
gtatgtgagtctattgctaaattggaagacgctactactattattggtggtggtgactct
gccgcagcagcaatttcattaggtttcgaagatgactttacacacatctctactggtgga
ggagcatcattagaatacttagaaggtaaagaattacctggtataaaagctattaatgat
aaɔtɔ a
```

Sequence 2216
MEKEIKFIGGVVNDPQKPVVAILGGAKVSDKINVIKNLVNIADKILIGGGMAYTFIKAQG
KEIGLSLLEEDKIDFAKDLLENNGDQIVLPVDCKIAKEFSNDAKITEVSINEIPSDQEAM
DIGPKTVELFNKELQGAHTVVWNGPMGVFEFSNFAKGTIGVCESIAKLEDATTIIGGGDS
AAAAISLGFEDDFTHISTGGGASLEYLEGKELPGIKAINDK*

Sequence 2217
Contig_0728_pos_6127_5366,
is similar to (with p-value 2.0e-90)
>sp:sp|P35144|TPIS_BACME TRIOSEPHOSPHATE ISOMERASE (EC 5.3.1
   .1) (TIM). >pir:pir|JQ1955|JQ1955 triose-phosphate isomerase
   (EC 5.3.1.1) - Bacillus megaterium >gp:gp|M87647|BACPGKTIMG
   _4 Bacillus megaterium glyceraldehyde-3-phosphate dehydrogen
   ase (gap), phosphoglycerate kinase (pgk), and triose phospha
   te isomerase (tpi) genes, complete cds. NID: g143315. >gp:gp
   |M87648|BACTPIPGK_2 Bacillus megaterium triose phosphate iso
   merase (tpi) gene, complete cds, and phosphoglycerate kinase
   (pgk) gene, 3' end. NID: g143759.

```
atgagaaɔaccaattatagccggaaactggaaaatgaataaaacagttcaagaagɔtaaa
gactttgtaaacgaattaccaacattacctgatcctaaagaagtagaatcagttatttgt
gcaccaacaatccaattagacgctttagtaacagctgttaaagatggtaaagcaaaaggg
ttaaaaattggagcacaaaacgcttactttgaagaaagcggtgcttatactggagaaact
tcaccagtagcattatctgaattaggtgttaaatatgtagtgattggtcactcagagcgt
cgtgactatttccacgaaactgacgaagaagtaaacaaaaaagcgcatgctatcttcaat
cacggtatgacacctattatttgtgtaggtgaatctgatgaagaacgtgaagctggtaaa
gcaaataaaatcgtaggtaatcaagtgaaaaaagctgtcgaaggtttatcagatgatcaa
cttaaagaagttgttattgcatatgaaccaatttgggctatcggtactggtaagtcatct
acatctgaagatgcaaatgaaatgtgtgctcacgtacgtcaaacattagctgacttatct
agtcaagaggttgctgacgctacacgtattcaatatggtggtagtgttaaacctaataac
attaaagaatatatggctcaatcagatatcgatggcgctcttgtaggtggcgcatcatta
aaagttgaagatttcgtacaattgttagaaggtgcaaaataa
```

Sequence 2218
MRTPIIAGNWKMNKTVQEAKDFVNELPTLPDPKEVESVICAPTIQLDALVTAVKDGKAKG
LKIGAQNAYFEESGAYTGETSPVALSELGVKYVVIGHSERRDYFHETDEEVNKKAHAIFN
HGMTPIICVGESDEEREAGKANKIVGNQVKKAVEGLSDDQLKEVVIAYEPIWAIGTGKSS
TSEDANEMCAHVRQTLADLSSQEVADATRIQYGGSVKPNNIKEYMAQSDIDGALVGGASL
KVEDFVQLLEGAK*

Sequence 2219
Contig_0728_pos_5363_3846,
is similar to (with p-value 0.0e+00)
>sp:sp|P39773|PMGI_BACSU 2,3-BISPHOSPHOGLYCERATE-INDEPENDENT
   PHOSPHOGLYCERATE MUTASE (EC 5.4.2.1) (PHOSPHOGLYCEROMUTASE)
   (BPG-INDEPENDENT PGAM) (VEGETATIVE PROTEIN 107) (VEG107).
```
atggcaaacaaccaactgccttaatcatcttagatggtttcgcaaatcgtgaaagtgaa
catggcaatgcagttaagcaagcacataaacctaattttgatcgatattatgaaaaatat
cctacaacacaaatagaagctagtggcttagatgtaggtcttcctgaaggtcaaatgggt
aactctgaagtaggacatatgaatattggtgcaggacgcatcgtatatcaaagtttaact
```

cgtaktaataaatcgattgaagacggagaattctttgataacactgtattaaatacgct
gttaaacatgttaaagacaatggctctgcgcttcatgtattcggattgctttctgatggt
ggtgtacacagtcattataagcatctatttgctattttagaattagctaaaaagcaagga
atagataaagtatatgtccacgcatttttagatggtcgtgatgttgatcaaaaatctgct
5   ttgaaatatatagaggaaactgaagataaatttaaagaattaggtgtaggccaattcgct
tctgtttcaggacgttattatgctatggaccgtgacaagcgttgggatcgtgaggaacgt
gcctataatgctattcgtaactttgaaggtcctacatttacttcagctaaagcaggcgtt
gaagctaattataaaaatgatgtgactgatgaattcgtcgaaccgtttatagttgaaggc
caaaacgatggtgtgaacgatggagacgcagtaatcttttataatttccgtccagataga
10   gcagctcaactttcagaaatctttactaataaagcgtttgatggatttaaagttgaacaa
gtggacaacttattctacgctacattcacgaaatataatgacaatgtagatgctgaaatt
gtatttgaaaaagttgacttaaataatacaatcggtgaagttgctcaagataatggcttg
aaacaattacgtatcgctgaaactgaaaagtatccacatgtaacatactttatgagtggt
ggacgaaatgaagagtttgaaggagaacgtcgtagactcatcgattctccaaaagtagcg
15   acttatgatttaaaacctgagatgagtgcatatgaagttaaagatgcattattagaagag
ttagacaaaggtgacttagatttaattctactgaactttgctaacccagatatggttgga
catagtggtatgcttgaaccaacaattaaagcaatcgaagcagtagatgagtgtcttggt
gaagtcgttgacaaaattattgatatgggtggtcatgccatcatcactgcagaccacggt
aactcagatcaagtattaactgatgacgaccaacctatgacgacacacaactaatcct
20   gttccagttattgtaactaaagaaggtgttacattaagagaaactggacgtttagacgat
ttagcgccgacattattagatttattaaatgttaaacaaccatctgaaatgacaggtgaa
tcactgattaaacattag

Sequence 2220
25   MAKQPTALIILDGFANRESEHGNAVKQAHKPNFDRYYEKYPTTQIEASGLDVGLPEGQMG
NSEVGHMNIGAGRIVYQSLTRINKSIEDGEFFDNTVLNNAVKHVKDNGSALHVFGLLSDG
GVHSHYKHLFAILELAKKQGIDKVYVHAFLDGRDVDQKSALKYIEETEDKFKELGVGQFA
SVSGRYYAMDRDKRWDREERAYNAIRNFEGPTFTSAKAGVEANYKNDVTDEFVEPFIVEG
QNDGVNDGDAVIFYNFRPDRAAQLSEIFTNKAFDGFKVEQVDNLFYATFTKYNDNVDAEI
30   VFEKVDLNNTIGEVAQDNGLKQLRIAETEKYPHVTYFMSGGRNEEFEGERRRLIDSPKVA
TYDLKPEMSAYEVKDALLEELDKGDLDLILLNFANPDMVGHSGMLEPTIKAIEAVDECLG
EVVDKIIDMGGHAIITADHGNSDQVLTDDDQPMTTHTTNPVPVIVTKEGVTLRETGRLGD
LAPTLLDLLNVKQPSEMTGESLIKH*

35   Sequence 2221
Contig_0728_pos_3707_2403,
is similar to (with p-value 0.0e+00)
>gp:gp|AF065394|AF065394_1 Staphylococcus aureus enolase (en
o) gene, complete cds. NID: g3152724.
40   atgccaattattacagatgtttacgctcgcgaagtcttagactcacgtggtaaccaaca
gttgaagttgaagtattaactgaaagtggtgctttcggacgtgcattagtaccttctggt
gcttctactggtgaacatgaagcagttgaattacgtgatggagatataaatcacgttattta
ggtaaaggtgtgactaaagcggtagaaaatgttaacgaaatgatcgcaccagaaatcgtt
gaaggtgaattttcagtttagatcaagtatctattgataaaatgatgattcaattagac
45   ggtacacacaacaaaggtaaattaggtgcaaatgccattttaggtgtttctattgccgta
gctcgtgcagctgctgacttattaggtcaaccattatataaatatttaggtggatttaat
ggtaaacaattgccagtacctatgatgaatattgttaatggtggttctcactcagatgca
ccaattgctttccaagagttcatgattttacctgtaggtgctgagtcattcaaagaatca
ttacgttggggtgcagaaatcttccataaccttaaatcaatcttaagtgaacgcggttta
50   gaaactgcagtaggtgatgaaggtggtttcgcacctagatttgaaggcactgaagacgct
gtagaaactattattaaagctatcgaaaaagcaggatacaaaccaggtgaagatgtattc
ttaggatttgactgtgcttcttctgaattctatgaaatggtgtttatgattacactaaa
ttcgaaggtgaacacggtgctaaacgtagtgcagcagagcaagttgactacttagaagaa
ttaattggtaaatatccaatcatcactattgaagatggtatggatgaaaacgattgggaa
55   ggttggaaacaattaactgatcgtatcggtgataaagttcaattagttggtgatgattta
ttcgtaactaacactgaaattttatctaaaggtatcgaacaaggtattggtaactcaatc
ttaatcaaagtaaaccaaatcggtacattaactgaaacattcgatgctattgaaatggct
caaaaagctggatatactgcggttgtatctcaccgttctggtgaaactgaagatactaca
attgctgatatcgcagttgctacaaatgcaggccaaattaaaacaggttcattatctaga

actgaccgtattgctaaatacaatcaattattacgtattgaagatgaattatacgaaaca
gctaaatttgaaggaattaaatctttctacaatttagataaataa

Sequence 2222
MPIITDVYAREVLDSRGNPTVEVEVLTESGAFGRALVPSGASTGEHEAVELRDGDKSRYL
GKGVTKAVENVNEMIAPEIVEGEFSVLDQVSIDKMMIQLDGTHNKGKLGANAILGVSIAV
ARAAADLLGQPLYKYLGGFNGKQLPVPMMNIVNGGSHSDAPIAFQEFMILPVGAESFKES
LRWGAEIFHNLKSILSERGLETAVGDEGGFAPRFEGTEDAVETIIKAIEKAGYKPGEDVF
LGFDCASSEFYENGVYDYTKFEGEHGAKRSAAEQVDYLEELIGKYPIITIEDGMDENDWE
GWKQLTDRIGDKVQLVGDDLFVTNTEILSKGIEQGIGNSILIKVNQIGTLTETFDAIEMA
QKAGYTAVVSHRSGETEDTTIADIAVATNAGQIKTGSLSRTDRIAKYNQLLRIEDELYET
AKFEGIKSFYNLDK*

Sequence 2223
Contig_0728_pos_2237_1779,
putative peptide of unknown function
atgacagactcaaatgctaaagaaataagaactggacgtttaattgcgataagttcatta
gtgttttgtattttacttatcatacaccactttattgtattagatgaatcaacagctaaa
tcaattttatctttagctggtcaaaaaacatcagatacagcagtgaaaaacatttaaat
agtgaccgatacactggaattatgtatattttagcttacttagcaggtactgttgctttc
tggaatcgccatccatatttatggtggtttatgtttgccgtatatatttctaatgcacta
tttacactcgtaaatctttacttatttattcaaggtattttagatgtaaaaaatgtactt
gcagttttaccaattttaattgtagtgattggatctataattctagcaatttatatgcta
gttgtttctattacacgtaaaagtactttcaatagatag

Sequence 2224
MTDSNAKEIRTGRLIAISSLVFCILLIIHHFIVLDESTAKSILSLAGQKTSDTAVKNILN
SDRYTGIMYILAYLAGTVAFWNRHPYLWWFMFAVYISNALFTLVNLYLFIQGILDVKNVL
AVLPILIVVIGSIILAIYMLVVSITRKSTFNR*

Sequence 2225
Contig_0728_pos_1710_1147,
is similar to (with p-value 3.0e-60)
>pir:pir|A40585|A40585 DNA topoisomerase (ATP-hydrolyzing) (
EC 5.99.1.3) chain B - Staphylococcus aureus >gp:gp|X71437|S
AGYRREC_2 S.aureus genes gyrB, gyrA and recF (partial). NID:
 g296393. >gp:gp|D10489|STAGYRABA_1 Staphylococcus aureus ge
nes for DNA gyrase A and B, complete cds. NID: g540540.
atgcatacattaatcatcgttttattaattatagattgtattgcattagtgactgttgta
ttactccaagaaggtaaaagtaatggactttcaggtgctattagtggtggcgctgaacaa
ttgtttggtaaacaaaaacaacgtggcgtcgatttattcttgcatagattaacaatacgt
acgttattacttacattcttctatcgtttcatgagacctttaattgaagcgggctacgtt
tatattgctcagccgcctttatataaactaacacaaggaaacaaaaatattatgtattt
aacgatagagaactagacaagttgaaacaagaattaaacccgtcaccaaaatggtcaatt
gcacgttacaaaggtcttggtgaaatgaacgcagaccaattatgggaaacgactatgaat
cctgaacatcgctctatgttgcaagtgagacttgaagatgcaattgatgcagaccaaaca
tttgaaatgttaatgggcgatgtagtagaaatcgcagacaatttatcgaagacaatgca
gtttatgccaacctagatttctag

Sequence 2226
MHTLTIVLLIIDCIALVTVVLLQEGKSNGLSGAISGGAEQLFGKQKQRGVDLFLHFLTIR
TLLLTFFYRFMRPLIEAGYVYIAQPPLYKLTQGKQKYYVFNDRELDKLKQELNPSFKWSI
ARYKGLGEMNADQLWETTMNPEHRSMLQVRLEDAIDADQTFEMLMGDVVENRRQFIEDNA
VYANLDF*

Sequence 2227
Contig_0728_pos_0_1110,
is similar to (with p-value 0.0e+00)
>sp:sp|P20831|GYRA_STAAU DNA GYRASE SUBUNIT A (EC 5.99.1.3).

```
     atggctgaattacctcaatcaagaattaatgaacgaaatataaccagtgaaatgcgtgaa
     tcattcttagactatgctatgagtgttatcgttctcgtgcattacctgatgttagagac
     ggattaaagccagtacatcgtcgtattctttatggtttaaatgaacaaggtatgacgccc
     gataaaccttataagaaatctgcacgtatagtcggggatgtcatgggtaaatatcaccct
 5   catggtgattcttcaatttatgaagcaatggtaagaatggcccaagactttagttatcgt
     tatccacttgtagatggtcaaggtaactttggctctatggatggtgacggtgcagctgca
     atgcgttataccgaagcacgtatgactaaaataacattagaactttacgtgatattaac
     aaagacacaattgattttattgacaactatgatggtaatgaaagagagccgtcagtctta
     cctgcacgtttccctaacttactagtaaatggtgcggcaggaattgccgtaggtatggct
 10  acaaatattcctccccacaatttaactgaagttattgatggtgtgctcagtttaagtaag
     aatccagacatcacaattaatgagctgatggaagacatacaaggtcctgattttcctaca
     gctggtttagtactagggaaaagtggtattcgtcgagcttatgaaacaggtcgtgggtca
     attcaaatgcgttctcgtgctgaaatagaagaacgtggtggtggccgtcaacgtattgtc
     gtaacggaaatacctttccaagtcaataaagcgcgtatgattgaaaaaatcgcagagtta
 15  gttagagataagaaaatcgacggtattacagatttacgtgatgaaacaagtttgcgtaca
     ggtgtaagagtagttattgatgtacgtaaagatgcaaatgcgagtgttattttaaataat
     ttatataaacaaacgccattacaaacatcatttggtgtgaatatgattgctttagtgaat
     ggtagacctaaactaatcaatttaaaagaagcacttatccattacttagaacaccaaaaa
     acagtggttagacgacgtactgaatataat
 20
     Sequence 2228
     MAELPQSRINERNITSEMRESFLDYAMSVIVSRALPDVRDGLKPVHRRILYGLNEQGMTP
     DKPYKKSARIVGDVMGKYHPHGDSSIYEAMVRMAQDFSYRYPLVDGQGNFGSMDGDGAAA
     MRYTEARMTKITLELLRDINKDTIDFIDNYDGNEREPSVLPARFPNLLVNGAAGIAVGMA
 25  TNIPPHNLTEVIDGVLSLSKNPDITINELMEDIQGPDFPTAGLVLGKSGIRRAYETGRGS
     IQMRSRAEIEERGGGRQRIVVTEIPFQVNKARMIEKIAELVRDKKIDGITDLRDETSLRT
     GVRVVIDVRKDANASVILNNLYKQTPLQTSFGVNMIALVNGRPKLINLKEALIHYLEHQK
     TVVRRRTEYN


 30  Sequence 2229
     Contig_0730_pos_6751_0,
     putative peptide of unknown function
     atgcttaattctacaatcaaatatgacactaagaagaaaaaactacctaaatttagtaaa
     ggtactaaaaagaaagacggtatattagatgttattagctctggtgtaaaaaatgatgtt
 35  aataaagtaaaagacattggtggtaaagcaagagacataggtggtactacgtttgacaaa
     gcaaaagacataggtacaaaagcacttgataaagctaaagatgtgtctagcactgttatc
     aagggtattggagatgtttttgattatgtaggtcatcctatgaaattggtaaataaagtc
     tttgagaaagttggttttaacctagactttatgaaaaatgcaccattaccatttgattta
     atgacagctatgattaagaaacttaaaaatggtattaaagacttctttaatgaaggttta
 40  gactctgcaggcggtggagatggttcttcgttcactaaattcccaattactacggggtat
     tatcctaatggtggtgctcctggttatagttttaatggtggtgctcactttggtattgac
     tatggcgctccatatggtacaactatcaat


     Sequence 2230
 45  MLNSTIKYDTKKKKLPKFSKGTKKKDGILDVISSGVKNDVNKVKDIGGKARDIGGTTFDK
     AKDIGTKALDKAKDVSSTVIKGIGDVFDYVGHPMKLVNKVFEKVGFNLDFMKNAPLPFDL
     MTAMIKKLKNGIKDFFNEGLDSAGGGDGSSFTKFPITTGYYPNGGAPGYSFNGGAHFGID
     YGAPYGTTIN


 50  Sequence 2231
     Contig_0730_pos_6983_6669,
     putative peptide of unknown function
     gtgctagacacatctttagctttatcaagtgcttttgtacctatgtcttttgctttgtca
     aacgtagtaccacctatgtctcttgctttaccaccaatgtctttttacttattaacatca
 55  ttttttacaccagagctaataacatctaatataccgtctttcttttttagtacctttacta
     aatttaggtagttttttcttcttagtgtcatatttgattgtagaattaagcattatcaag
     gtcgcattcaatcaggtgagaagattaaatatactaaataatatattaggtgattcgtta
     tgcaaatattattag
```

Sequence 2232
VLDTSLALSSAFVPMSFALSNVVPPMSLALPPMSFTLLTSFFTPELITSNIPSFFLVPLL
NLGSFFFLVSYLIVELSIIKVAFNQVRRLNILNNILGDSLCKYY*

Sequence 2233
Contig_0730_pos_6435_6010,
is similar to (with p-value 3.0e-21)
>sp:sp|P42421|YXDJ_BACSU HYPOTHETICAL 26.6 KD SENSORY TRANSD
UCTION PROTEIN IN IDH 3'REGION. >gp:gp|D14399|BACIOLO_11 Bac
illus subtilis 15 kb chromosome segment contains the iol ope
ron. NID: g709980. >gp:gp|Z99124|BSUB0021_70 Bacillus subtil
is complete genome (section 21 of 21): from 3999281 to 42148
14. NID: g2636442.
atcgatcaagtgatgagtatggaacttggtgcagatgattatatgcaaaaaccatttat
acaaacgtcttaattgctaagctacaagctatttatagacgcgtttatgaatttgcagtt
gaagaaaagagaacgttaagttggcaagacgctactgtggatttatcaaaagatagtatt
caaaaagatgataaaactatctttttgtctaaaacagagatgattattttagagatgtta
atcaataaacgtaatcaaatcgtgacacgagacactctcattactgctttgtgggatgat
gaagctttttgttagtgataatactttaacagttaatgttaatagattaagaaaaaaatta
tcagaaattgacatggatagtgcaattgaaaccaaagttggtaaaggatacttagctcat
gaataa

Sequence 2234
MDQVMSMELGADDYMQKPFYTNVLIAKLQAIYRRVYEFGVEEKRTLSWQDATVDLSKDSI
QKDDKTIFLSKTEMIILEMLINKRNQIVTRDTLITALWDDEAFVSDNTLTVNVNRLRKKL
SEIDMDSAIETKVGKGYLAHE*

Sequence 2235
Contig_0730_pos_5606_4977,
putative peptide of unknown function
atgaaattattgatagatcaagagaatgatgatcagcgtaagcgagcgttattatttgaa
tggtctcgtattaatgagatgttagataagcaattatatttaacaaggcttgaaacacat
catcgtgatatgtattttgattatatttcattaaagagaatggttatagatgaaatacaa
gttactcgacatatcagtcaggcaaaagggataggtttttgaattagattttaaagacgaa
caaaagctttatacagatgttaaatggtgccgtatgatgattaggcaagttctatctaac
tctttgaaatatagtgataattctacaataaatttaagtggttataacatagaaggacac
gttgttttaaaaattaaagactacggtcgtggaattagtaaaagagatttaccacgtata
tttgatagaggatttacttctacaacagaccgcaacgatactgcgtcttctggtatggga
ttataccttgtacaaagcgtgaaagaacaacttgggattgaagttaaagttgattcaata
gtggggaaaggaacaacgtttatttcattttcccacaacaaaatgaaatcattgagcgc
atgtctaaagtgacaagattgtcattttaa

Sequence 2236
MKLLIDQENDDQRKRALLFEWSRINEMLDKQLYLTRLETHHRDMYFDYISLKRMVIDEIQ
VTRHISQAKGIGFELDFKDEQKVYTDVKWCRMMIRQVLSNSLKYSDNSTINLSGYNIEGH
VVLKIKDYGRGISKRDLPRIFDRGFTSTTDRNDTASSGMGLYLVQSVKEQLGIEVKVDSI
VGKGTTFYFIFPQQNEIIERMSKVTRLSF*

Sequence 2237
Contig_0730_pos_4513_4070,
is similar to (with p-value 8.0e-32)
>sp:sp|P42423|YXDL_BACSU HYPOTHETICAL ABC TRANSPORTER ATP-BI
NDING PROTEIN IN IDH 3'REGION. >gp:gp|D14399|BACIOLO_13 Baci
llus subtilis 15 kb chromosome segment contains the iol oper
on. NID: g709980. >gp:gp|Z99124|BSUB0021_68 Bacillus subtili
s complete genome (section 21 of 21): from 3999281 to 421481
4. NID: g2636442. >gp:gp|D45912|D45912_2 Bacillus subtilis g
enome sequence between the iol and hut operon, partial and c
omplete cds. NID: g1408482.

atgcttccactatcagttcaaaagttagataaacaaataatgcatgaacgttatcaacgt
atagtagaagctttgaacattagtgatattagtgataaatatccatcagagttgtcaggt
ggacagcgtcaacgtacctcagcagcaagggcatttattaatttaccttcaattatatt
gctgatgagcctacaggtgctttagattctaaaagtacactagatttacttaagcgtctt

5   aaatatatgaatgaggaatttaacacaaccatacttatggtgacacatgatccagtagcg
gctagtttttcaaaccgtgttgtgatgttaaaggatggacaaatctttactgagttgtat
caaggtgatgatgataaacaaacgttttataaagaaattataagaacgcaaagtgtactt
ggtggcatcaattatgagctttaa

10   Sequence 2238
MLPLSVQKLDKQIMHERYQRIVEALNISDISDKYPSELSGGQRQRTSAARAFINLPSIIF
ADEPTGALDSKSTLDLLKRLKYMNEEFNTTILMVTHDPVAASFSNRVVMLKDGQIFTELY
QGDDDKQTFYKEIIRTQSVLGGINYEL*

15   Sequence 2239
Contig_0730_pos_3876_3457,
putative peptide of unknown function
gtggtctttttgctttatgctaattttttattttttgaaaagacgaggtcgtgaactatcg
ttattacaaattattggtctaacaaagaaagatatcatgaaaatgattatgttggagcaa

20   ttgatgacatttatgatgacaactattgtaggtatcatattgggaatctttggttcgaaa
attttactcatgattgtattgcgattattaggaatcaacgtgagtgtttctattatattt
aattatcatgccatttttagaaacgttattattaatagctgtgtcatatgtacttatagtc
tttcaaagctatgtatatttacttaaacgttctattaaagagttagcgtctgatgtaaat
aaaaaagagttcagtcatacacgcacaacacttggtgaagttgtattaggtttcttataa

25

Sequence 2240
VVFLLYANFLFLKRRGRELSLLQIIGLTKKDIMKMIMLEQLMTFMMTTIVGIILGIFGSK
ILLMIVLRLLGINVSVSIIFNYHAILETLLLIAVSYVLIVFQSYVYLLKRSIKELASDVN

30   KKEFSHTRTTLGEVVLGFL*

Sequence 2241
Contig_0730_pos_3418_2012,
is similar to (with p-value 2.0e-78)

35   >sp:sp|Q02001|TRPE_LACLA ANTHRANILATE SYNTHASE COMPONENT I (
EC 4.1.3.27). >pir:pir|S35124|S35124 anthranilate synthase (
EC 4.1.3.27) alpha chain - Lactococcus lactis subsp. lactis
>gp:gp|M87483|LACTRPOP_2 L. lactis trpE, trpG, trpD, trpF, t
rpC, trpB trpA genes, complete cds. NID: g149514.

40   atggatattgtatacaaaaaggtgaatgctcaaattacgccagaagctttagcaaaatta
aaacaaaaaaagatcatttttgaaagtacaaatcaacagaaacttaaaggtaggtactcg
atagtagtattcgatcattatggcaaaattacattagataattctcaacttttaattaag
ttagacaatcattgtgaaatagttaagaatcaaccgtatcaacgacttaaggaatttgta
gataaatattattttgaaatcaaagataaatatttaaaagatttaccttttatttcgggc

45   tttatagggacatgtagctttgatttagtacgacatgaatttaaaaaattacaagatatt
aaattagaagatcatcaaactcatgatgtccaattttatctagtggaagatgtatttgtt
tttgatcattataaagatgaattatatattatcgcaagtaacttattttcttatagaaca
aaagagagattaaaggaatctattgaacgtaaaattgaagatttaaaaaacatacatttt
tcggttgaggatataaattataaatccatccctcgacatataaccaccaatatatcagag

50   caacaatttgttcaaactattagaattttaaaaaagaaaattactgaaggagatatgttt
caagtagttccttcaagaatttatagttataaacaccattttcaacacaatttacatcaa
ttaacttttcagttatatcaaaatttaaagcgacaaaatcctagtccatatatgtattat
attaataaagatgtaccgattgtaataggaagttctcctgaaagttttgtaaaggtaaaa
gatggaaaagtttatacgaatcctatagctggaacaattaaaagaggtcaaaataaaaaa

55   gaagatgaaaataatgaaaagacattaatgaaagatgaaaaggaattgagtgaacatcgt
atgctcgtagatttaggaagaaatgatattcatcgaataagtaaaacaggcacttcacaa
attaccaaactaatgacaatagaacgttatgaacatgtcatgcatatcgttagtgaagtt
attggagaattaaaaccccatctatctcctatgagcgtcatcgcaagtttgctaccaacg
ggtactgtctcaggtgcacctaaacttagagctatacagagaatatacgaatcttatcct

tataaaagaggtatctatagcggtggtgttgggtatatcaactgtaatcatcatttagat
tttgcattggctatacgtaccatgattatcgatgaggaaaaagtcagtgtcgaggcagga
tgtggagtagtatatgattctattccagagaaagaacttgaagaaacaaaacttaaagct
aaaagtttattggaggtaactccatga

5

Sequence 2242
MDIVYKKVNAQITPEALAKLKQKKIIFESTNQQKLKGRYSIVVFDHYGKITLDNSQLLIK
LDNHCEIVKNQPYQRLKEFVDKYYFEIKDKYLKDLPFISGFIGTCSFDLVRHEFKKLQDI
KLEDHQTHDVQFYLVEDVFVFDHYKDELYIIASNLFSYRTKERLKESIERKIEDLKNIHF

10   SVEDINYKSIPRHITTNISEQQFVQTIRILKKKITEGDMFQVVPSRIYSYKHHFQHNLHQ
LTFQLYQNLKRQNPSPYMYYINKDVPIVIGSSPESFVKVKDGKVYTNPIAGTIKRGQNKK
EDENNEKTLMKDEKELSEHRMLVDLGRNDIHRISKTGTSQITKLMTIERYEHVMHIVSEV
IGELKPHLSPMSVIASLLPTGTVSGAPKLRAIQRIYESYPYKRGIYSGGVGYINCNHHLD
FALAIRTMIIDEEKVSVEAGCGVVYDSIPEKELEETKLKAKSLLEVTP*

15

Sequence 2243
Contig_0730_pos_1446_451,
is similar to (with p-value 8.0e-51)
>sp:sp|P17170|TRPD_LACCA ANTHRANILATE PHOSPHORIBOSYLTRANSFER

20   ASE (EC 2.4.2.18). >pir:pir|S42343|JS0340 anthranilate phosp
horibosyltransferase (EC 2.4.2.18) - Lactobacillus casei >gp
:gp|D00496|LBATRP_2 Lactobacillus casei DNA, trp operon (trp
D, trpC, trpF, trpB, trpA), complete cds. NID: g216754.
atgacccttcttgagaaaattaaacaaataaatctttatctaaaaaagatatgcaatca

25   tttattgttacactgtttgattcaaatatagaaaccaatgtaaaggttgaattattgaaa
gcttatacaaataaagacatgggtcaatatgagctaacgtatttagttgaatattttatc
cagacaaactatccaaaccaaccattttataataaagctatgtgtgtttgtggcacaggt
ggagatcaatcaaatagctttaatatttctacaactgtagcttttgttgtagcaagtgca
ggagtgccagtcattaaacacggtaataaaagtattacttcacattcaggaagtacagat

30   gtattacatgaaatgaatataaaaacaaacaaaatgaacgaagtagagcaacaattaaat
ttgaaaggattagcattcataagtgcaactgattcttatccaatgatgaaaaagcttcaa
tcaattagaaaatcgattgcaacacctacaatttttaacttgattggaccattaattaat
cctttcaaattaacttatcaagtgatgggggtatatgaagcttcacaacttgaaaatata
gcacaaacattaaaggatttaggtagaaaacgagcaattttaattcatggtgcaaatggg

35   atggatgaggccacgctttctggtgaaaatatcatttatgaagttagcagcgaaagagca
ttaaaaaaatatagtttaaaagcagaagaagtcggttttagcttatgcaaataatgacacg
ttgatagytggttcacctcaaacaaataaacaaattgcattgaatatcctaagtggcacg
gatcactcaagtaaacgagatgtagttttgttaaatgctggaattgctttatatgttgct
gagcaagtggaaagtatcaaacatggcgtagagagagcgaaatatctcattgatacaggt

40   atggcaatgaaacaatatttaaaaatgggaggttaa

Sequence 2244
MTLLEKIKQNKSLSKKDMQSFIVTLFDSNIETNVKVELLKAYTNKDMGQYELTYLVEYFI
QTNYPNQPFYNKAMCVCGTGGDQSNSFNISTTVAFVVASAGVPVIKHGNKSITSHSGSTD

45   VLHEMNIKTNKMNEVEQQLNLKGLAFISATDSYPMMKKLQSIRKSIATPTIFNLIGPLIN
PFKLTYQVMGVYEASQLENIAQTLKDLGRKRAILIHGANGMDEATLSGENIIYEVSSERA
LKKYSLKAEEVGLAYANNDTLIGGSPQTNKQIALNILSGTDHSSKRDVVLLNAGIALYVA
EQVESIKHGVERAKYLIDTGMAMKQYLKMGG*

50   Sequence 2245
Contig_0730_pos_447_70,
putative peptide of unknown function
atgactattttaaatgaaattattgagtataaaaaaactttgcttgagcgtaaatactat
gataaaaaacttgaaattttacaagataacggaaatgttaagaggagaaagctgattgat

55   tcacttaactatgatagaacattatcagttattgctgaaataaaatcgaaaagcccatct
gtacctcaattaccgcaacgtgatcttgttcaacaagttaaagattatcaaaaatatggt
gctaatgctatttcaatattaactgatgaaaaatactttggcggtagttttgaacgatta
aatcagttatcaaagataactgaacgtttaacctttttgtacctcatcagcttgctcttcg
acgcccatattcttatga

Sequence 2246
MTILNEIIEYKKTLLERKYYDKKLEILQDNGNVKRRKLIDSLNYDRTLSVIAEIKSKSPS
VPQLPQRDLVQQVKDYQKYGANAISILTDEKYFGGSFERLNQLSKITERLTFCTSSACSS
TPIFL*

Sequence 2247
Contig_0732_pos_871_1170,
putative peptide of unknown function
atggaccatattccacaagtagttattttaataaaaaagacttatgtaacgaacagatg
gatgtacctgtatctaaatctgcgcatgttttgtatctagtcgtgatgaaatgataaa
caaaaggtgaaaaatttagtaattcaagaaataaaaaatagtctcagcccatacgaagaa
attgtagatagtgctgatgcagatagattatattttcttaaacaacacacgcttgttact
gaattaatatttgacgaaacacaagcatcttatcgtatcaaaggatttaaaaaattataa

Sequence 2248
MDHIPQVVIFNKKDLCNEQMDVPVSKSAHVFVSSRDENDKQKVKNLVIQEIKNSLSPYEE
IVDSADADRLYFLKQHTLVTELIFDETQASYRIKGFKKL*

Sequence 2249
Contig_0732_pos_3289_4629,
is similar to (with p-value 0.0e+00)
>sp:sp|Q59812|GLNA_STAAU GLUTAMINE SYNTHETASE (EC 6.3.1.2) (
GLUTAMATE--AMMONIA LIGASE) (GS). >gp:gp|X76490|SAGLNAR_3 S.a
ureus (bb270) glnA and glnR genes. NID: g1134885.
atgccaaaacgtagttttacaaaagatgatattcgtaaatttgctgaagaagaaacgta
agatatttaagattacaattcactgatattttagggactattaaaaatgttgaagttcca
gtaactcaattagaaaaagtattagataatgaaatgatgtttgatggttcatctatgaa
ggtttcgttcgtatcgaagaatcagatatgtatttacatcctgatttagatacttgggtt
atcttcccttggactgctggacaaggaaaagttgcacgactaatctgtgatgtatttaaa
acagatggtacaccatttgaaggtgatccacgagctaacttgaagcgtgtattaagaaga
atggaagatatgggctttactgattttaatctagggcctgaaccagaattttcttattt
aaattagacgaaaaaggcgaacctacattagaattaaacgatgatggtggttatttcgat
ttagctcctacagatttaggtgaaaattgtcgccgtgacatcgtttttagaattagaagat
atgggctttgacattgaagcaagccaccatgaagtagcgccaggtcaacatgaaattgac
tttaaatatgcagatgccgttacagcatgtgataatatccaaacatttaaactagttgtt
aaaacaattgcacgtaagcataatttacatgcaacatttatgccaaaaccattatttggt
gtaaacggtagtggtatgcacttcaacgtatcactatttaaaggaaaagagaatgcgttc
tttgatcctgaaggtgatttacaattgactgatactgcatatcaatttacagctggtgtc
cttaaaaacgctagaggattcactgcagtatgtaatccaattgtcaactcatataaacgt
cttgtaccaggttacgaagcaccatgttatattgcatggagtggtaaaaaccgttcacct
ttagtacgtgttccaacatctagaggtctatcaactcgtattgaagtacgctcagttgac
cctgcagctaacccgtacatggcattagcagcaatcttagaagcagggttagatggaatt
gagaataaacttgaggttccagaacctgtaaaccaaaatatctacgaaatgaatcgtgaa
gaacgagaagcggttggtatccaagacttaccttcaactttatacactgcgttaaaagca
atgcgtgaaaataaatcaattaaaaacgcattaggtaatcatatttacaatcaatttatt
aacTcaaratcgattgaatgggattactatagaactcaagtatccgaatgggaacagaa
cagtatactaagcaatactaa

Sequence 2250
MPKRSFTKDDIRKFAEEENVRYLRLQFTDILGTIKNVEVPVSQLEKVLDNEMMFDGSSIE
GFVRIEESDMYLHPDLDTWVIFPWTAGQGKVARLICDVFKTDGTPFEGDPRANLKRVLRR
MEDMGFTDFNLGPEPEFFLFKLDEKGEPTLELNDDGGYFDLAPTDLGENCRRDIVLELED
MGFDIEASHHEVAPGQHEIDFKYADAVTACDNIQTFKLVVKTIARKHNLHATFMPKPLFG
VNGSGMHFNVSLFKGKENAFFDPEGDLQLTDTAYQFTAGVLKNARGFTAVCNPIVNSYKR
LVPGYEAPCYIAWSGKNRSPLVRVPTSRGLSTRIEVRSVDPAANPYMALAAILEAGLDGI
ENKLEVPEPVNQNIYEMNREEREAVGIQDLPSTLYTALKAMRENKSIKNALGNHIYNQFI
NSKSIEWDYYRTQVSEWEREQYIKQY*

Sequence 2251
Contig_0732_pos_6496_6840,
putative peptide of unknown function
atgaataaagaccaaaagaatcaatatgagttggaaaagctgttaaaagaaaatgaagaa
ctaaaagcagaaaaagctttatctcaaatgaagaatgagactcgttcaatgcttaatgag
tcaggtttagaaaacttcgatgatcaaattgttaatatattagtaaatactgatgctgaa
aaaacaaggaaaaatgttgaatcatttactaacttacttaatcaaatggtaaaatcaaat
gttgaaaaagcattaagacaagactcaccagtaagcactcaatcaaataaaatgacaaaa
gatgaagaatcagattatcttgtttcagggagaaaatgtctataa

Sequence 2252
MNKDQKNQYELEKLLKENEELKAEKALSQMKNETRSMLNESGLENFDDQIVNILVNTDAE
KTRKNVESFTNLLNQMVKSNVEKALRQDSPVSTQSNKMTKDEESDYLVSGRKCL*

Sequence 2253
Contig_0732_pos_8090_8434,
putative peptide of unknown function
atgctaattattttaattttactagcattaagttttggtttaattcctatttctattaac
ttgcctatattaggtttaattccctttattttatggggagctatgggatgggctacacaa
gctcctcaacaacatatattattgaaaaaacatcctgaatatggaggctctgctgtcgct
ttaaatagttctattaattatttaggcagtgctatgggatcagcaatcggaggaattatt
ttatttaatgctaatagtacaaatgtactaatatatagtgctttaggaattactattatt
ggtattttattacaattactaaatttatccctagaaaaaaattaa

Sequence 2254
MLIILILLALSFGLIPISINLPILGLIPFILWGAMGWATQAPQQHILLKKHPEYGGSAVA
LNSSINYLGSAMGSAIGGIILFNANSTNVLIYSALGITIIGILLQLLNLSLEKN*

Sequence 2255
Contig_0732_pos_7495_7091,
putative peptide of unknown function
atggaaaaatcaacgcagacagacaaacaaaactctgtgaacttaaagcaaaacacaaaa
gatcaaaataataacgcaaatgatgaagcagcttctccaactagcgaacaaaatgcagct
atagcacaagcaaagtcatatgcaaatacattacctatctctaagaaaagtttatacaaa
caattaacttcggaatacggagagaaatatccggcagacatagcacagtatgctgttgac
catatcagtgtagattataaaatgaatgcactgagattagcaaaaagttacgtaaaaaat
ataaacatttctaatcaagcgttatatgatcaactcgtttcagaaaatggagaaggattt
actcctgaagaagcacaatatgcaatgaatcatttagataggtaa

Sequence 2256
MEKSTQTDKQNSVNLKQNTKDQNNNANDEAASPTSEQNAAIAQAKSYANTLPISKKSLYK
QLTSEYGEKYPADIAQYAVDHISVDYKMNALRLAKSYVKNINISNQALYDQLVSENGEGF
TPEEAQYAMNHLDR*

Sequence 2257
Contig_0733_pos_5081_5728,
putative peptide of unknown function
gtgtctacttcccaattgattgtttcgaattccggacgagctaactcagggtttttctct
aattcagcaacagtgttgaatttagcgttagcacgttttaagatcgggtatttcccactt
gcagttgatactgaagttttttgtaccaattctgataagtcttggactgtcttaacttct
ttttcaggaatatatttaatatcctctgggatagttacgccaacgtcatcagatttaaca
ttgtcacgtttagccccttttgatttcatgtactgttcaaatgctagaatttcttcgttt
gtctctggattttggtttaatttagccatagaacgtttcgctccttctttttttgtctttt
tctttttttaattcttcttctgttggttcttctacttttttcaatagtaggtgtttctggt
gtttcttcaggtttgtcatctggttttggtgcatcatcaggttttttcttcatctgaagtt
ccttctggttcatcatcagaaggtttgttctctgattcttctccagaattaccatctttg
ttatcttcaacttctgcaccttcatctttaggtggttcatcttgtttaggtgctgacgct
tcaatttcttttgaaagctgttcgagttcttcgtactctttcttttga

Sequence 2258
VSTSQLIVSNSGRANSGFFSNSATVLNLALARFKIGYFPLAVDTEVFCTNSDKSWTVLTS
FSGIYLISSGIVTPTSSDLTLSRLAPFDFMYCSNARISSFVSGFWFNLAIERFAPSFLSF
SFFNSSSVGSSTFSIVGVSGVSSGLSSGFGASSGFSSSEVPSGSSSEGLFSDSSPELPSL
LSSTSAPSSLGGSSCLGADASISFESCSSSSYSFF*

Sequence 2259
Contig_0733_pos_11496_10366,
putative peptide of unknown function
atgaattcgttgcacatagcaggacgtattttcaaacagacgattcgagatgtaagaaca
ttggcactgttacttattgcacctatattactattgtcgctactatattacatttttaca
gttgccgataatacgaatggcgtaacagttggggttcacgatgtaccagattcattaatg
actgaattacatgataaagatattcacgttaaacattataaaaatgacaatgatataagt
gataaaattaaagacgacaaattaacaggattttttgcacagtgatggtcaaaaagtatca
gtgacttatgctaacgataatcctacacaagcaggagaactaacaggtgcaaatcaaaaa
tggttaatgagtcataacatgaatgccatgaaagataatactaataaattgcatcaagcg
ttaactaaaatacaacaaaaaatgcccgggatgggggagacacgcctcatcaagatatg
gctaaaccatataaactaacaacgcactatttatatggttcatcagattctacgtatttt
gatatgataaatcctattttaattggatttttttgtcttttttctttacgtttttaatttct
ggcattggcttattaaaagagcgtacttctggccattagaacgtttacttgcctctcca
ataaaaagaagtgaaattattttttggttatgttttcggttatggtagttttagcgtttatc
caaacaatagttgtcgtattatatgcaatttatattctgcatatagacttagtaggttcg
atatggttcgtactattaacggcaatattaacagcgcttgtcgctgtgacattcggtata
ttattatctacctttgcttcctcagaattccaaatgattcaatttataccattagtcata
gtgccacaagtactatttgcaggcattataccaattgaatcaatgaataaaggattacaa
tacttttcacatatcatgccgttattctataccggccaaacgatgcaaatattatgatc
aagggttatggattcaacgatatttacatttatttaattgtgttattcgcattttttcatt
ttcttattgattttaaatattataggcatgaaaagatatagaaaagtttag

Sequence 2260
MNSLHIAGRIFKQTIRDVRTLALLLIAPILLLSLLYYIFTVADNTNGVTVGVHDV.DSLM
TELHDKDIHVKHYKNDNDISDKIKDDKLTGFLHSDGQKVSVTYANDNPTQAGELTGANQK
WLMSHNMNAMKDNTNKLHQALTKIQQKMPGDGGDTPHQDMAKPYKLTTHYLYGSSDSTYF
DMINPILIGFFVFFFTFLISGIGLLKERTSGTLERLLASPIKRSEIIFGYVFGYGSFSVI
QTIVVVLYAIYILHIDLVGSIWFVLLTAILTALVAVTFGILLSTFASSEFQMIQFIPLVI
VPQVLFAGIIPIESMNKGLQYFSHIMPLFYTGQTMQNIMIKGYGFNDIYIYLIVLFAFFI
FLLILNIIGMKRYRKV*

Sequence 2261
Contig_0733_pos_10356_9706,
putative peptide of unknown function
atgaaccaagatattaagtcattagttgaaaccattgtgcctcaacttgaatatttaagc
gataaacaaagacgtgtcatagaaagtgctattgcattattcagtgaacaaggatttgat
aaaacgagtactaaagaaattgcgcagcgtgcaaatgtcgcagaaggaacggtatttaag
cagtttaaaagtaaaagaatgttattatacgcaggattaattccaattttaagagatcat
atcgcacctgtagctgttaaacaatttacagatgaattaaacgaagtaacccatttttgat
gcatttataaatttatttgtagaaaatagatctaaatttatttatgacaatagacgtatt
cttaaagtcatcttaaatgaagctattactaatgaagattttcaaaatatattagttaat
attttcacccataaaattaacgagtaaattaaaagataaaattgaatggtttatcgataat
ggtggcatgcgcaatgttaaacctgagttttttatacgtacggtcgtcgcacaaatctta
aatttaaatatcccaataatagttaataatgactatactaagggtgaaaactatcagcag
tttgcgttattcgtaaaagagggcttatataggatgtttaagcgagaatag

Sequence 2262
MNQDIKSLVETIVPQLEYLSDKQRRVIESAIALFSEQGFDKTSTKEIAQRANVAEGTVFK
QFKSKRMLLYAGLIPILRDHIAPVAVKQFTDELNEVTHFDAFINLFVENRSKFIYDNRRI
LKVILNEAITNEDFQNILVNIFTHKLTSKLKDKIEWFIDNGDMRNVKPEFFIRTVVAQIL
NLNIPIIVNNDYTKGENYQQFALFVKEGLYRMFKRE*

Sequence 2263
Contig_0733_pos_4489_3998,
putative peptide of unknown function
atgccacttttttgtgacagattatttaaatgaaaaggaaagatatgtaggctatttagca
agtttatgtgcaggtttagaagtgccgtttatggttatattaggtatattatcagctaaa
ttgccaactcgaactttattgataattggttctgtatttggtggcgcattttactttagt
attggtgtatttaaaaattttcatatgatgcttgttggacaaatatgtttagccatcttt
ctagcaatattactaggtcttgggattagttattttcaagatattcttcctgattttcca
ggttatgcatcaacacttttttgccaatgccatggtaataggacaacttggtggtaatttg
ctaggtggtgtgatgagtcattgggttggtttggaaaatgtattctttgtctcagcaagt
tctatcttcgtagggatgatactcatcttgtttacgaaagatcaaaaaattacaatagaa
gatgtggagtag

Sequence 2264
MPLFVTDYLNEKERYVGYLASLCAGLEVPFMVILGILSAKLPTRTLLIIGSVFGGAFYFS
IGVFKNFHMMLVGQICLAIFLAILLGLGISYFQDILPDFPGYASTLFANAMVIGQLGGNL
LGGVMSHWVGLENVFFVSASSIFVGMILILFTKDQKITIEDVE*

Sequence 2265
Contig_0733_pos_3991_3503,
putative peptide of unknown function
atgacagcgatattatggattttaattattatagcctttgcgttagcatttattggttta
attaaaccggtgataccatcactttttaatgttatggattggttttttaatttatcaattt
ggttttcatgagggaagattatcgtggatttttttatgttgcaatgattatctttaccatt
atgatattagtagccgattttgtgatgaataaatatttcgtcaatcgctttggaggaagt
aaaataggtgaatacacagcgctcataggtgtgattgtaggttgtttcgtttttcctccc
tttggtattatcattattcctttttgttgctgtgttcattgttgaattggttcaagggttt
aactttcaacaagctataaaggtgagtttttggctcagtgattgcatttttagcgagtaca
attgctcaaggtctaataatgattgtaatggttatttggttcttttttagatgtctttcta
ataaattaa

Sequence 2266
MTAILWILIIIAFALAFIGLIKPVIPSLLMLWIGFLIYQFGFHEGRLSWIFYVAMIIFTI
MILVADFVMNKYFVNRFGGSKIGEYTALIGVIVGCFVFPPFGIIIIPFVAVFIVELVQGF
NFQQAIKVSFGSVIAFLASTIAQGLIMIVMVIWFFLDVFLIN*

Sequence 2267
Contig_0733_pos_2319_1897,
putative peptide of unknown function
atgagaaaatggttaaccttactattaattacaacattggtgttaactgcatgtggtaaa
agtaacgaaaaagcttcttttagaaaaaagcattgatcagttgaaaaaagaaaataaggat
ttaacaaaacagaagaaaaagttacaagagcaaaaggataagcttaaacacaaacaggat
agtctccaagaagatgtaaatgacttgcctgctaaaagcacatcccgagataagaaaaat
aaagataatcatgatgcaaaagaaaagtcttcagataatcaatcgacatctgctaatcat
gatgatcaaactaacaaaataaaaagcaatcaagatgaacatgacagtcaatcctctaaa
ccacatacacagcagaagccctcacagaatgatagaaaaaataatcatcgacaagaacga
tag

Sequence 2268
MRKWLTLLLITTLVLTACGKSNEKASLEKSIDQLKKENKDLKKQKKKLQEQKDKLKHKQD
SLQEDVNDLPAKSTSRDKKNKDNHDAKEKSSDNQSTSANHDDQTNKIKSNQDEHDSQSSK
PHTQQKPSQNDRKNNHRQER*

Sequence 2269
Contig_0733_pos_765_244,
putative peptide of unknown function
atgtcaaaaatcttaaacacacaattaactggtattttttaatcggcttgaaaaacaagag
ttggatattcaaatggcagctcaatgtctcattcaagcaattggtggagaaggacatgtc

tatatcaaaggctacgatgatttaaaattctatgagtcattcatattacaaagccatgaa
aaattagcgtctagcttaccacttgaagatttacaaaattttaacgatatagatacaaca
gatagggtactgttattttcaccatactacacttcggaagttgaaagtgatgtacttcaa
cttattgatttagatgtcgatttagtgcttatttgtaataaccctaaacgagatgatttt
5   cctaatcatttaattcattatgttaatttatcaacacctaggcccattgtttacacagaa
gattatgataaaatcattcaaccacatccgatggccttaaattatatttattatgatatt
tatactcaaatgattgagatgactagagacctagatttatag

Sequence 2270
10   MSKILNTQLTGIFNRLEKQELDIQMAAQCLIQAIGGEGHVYIKGYDDLKFYESFILQSHE
KLASSLPLEDLQNFNDIDTTDRVLLFSPYYTSEVESDVLQLIDLDVDLVLICNNPKRDDF
PNHLIHYVNLSTPRPIVYTEDYDKIIQPHPMALNYIYYDIYTQMIEMTRDLDL*

Sequence 2271
15   Contig_0734_pos_2644_3024,
putative peptide of unknown function
atgcaccagcttgtccatactgacgataatgcaataaatgtagatgtattaccaccacaa
caagctgacggtaaagcgactaatccagaacaattatttgctgcaggttacgcttcatgc
tttaatggtgcatttgatttaattttaaaacaaataaagtgcgcgatgctgaacctgaa
20   gtaacgttaacggtacgcttggaagacgatccagatgccgaaagcccaaaacttagcgtt
gatattcatgcaaaagttaaaaatgttttatcacaagaagatgctgaaaaatatttacaa
gatgcgcacgactttgtccgtattcaaaagctacacgtggcaatatcgatgtaaactta
aatgttgaagtagtagaataa

25   Sequence 2272
MHQLVHTDDNAINVDVLPPQQADGKATNPEQLFAAGYASCFNGAFDLILKQNKVRDAEPE
VTLTVRLEDDPDAESPKLSVDIHAKVKNVLSQEDAEKYLQDAHDFCPYSKATRGNIDVNL
NVEVVE*

30   Sequence 2273
Contig_0734_pos_3557_4369,
is similar to (with p-value 1.0e-96)
>gp:gp|Y17554|BLY17554_1 Bacillus licheniformis arcA, arcB,
arcC and arcD genes. NID: g3687415.
35   atggatgataaatatccgtttttatcttgacccaatgcctaatctatatttcacacgtgat
cctcaagcttcaattggtagaggtatgacagtaaatcgtatgtttttggagagcgagacgc
agagaatcgattttcatttcatatattttaaaacatcatcctagatttaaagatgagaat
attcctttatgggtggatcgtgactgtccgttcaacatcgaaggtggagacgaactggtg
ttatctaaagatgtacttgcaatagggatatctgaacgtacttctgcacaagcaattgaa
40   cgtttagcacgacgtattttttaaagatccgttatctacttttaaaaaggtggtggcgatt
gagattccaactagtcgaacatttatgcacttagatactgtttgtacaatgattgattac
gacaaattcactacacattcagcaattcttaaatcagaaggaaacatgaatatctttatt
atcgaatatgatgataaagctgaagatatcaaaatccaacattctagtcatcttaaacaa
acattagaagaagtgctcgatgttgatgaaatcacattaataccaactggaaatggtgat
45   atcatcgacggtgctcgtgaacaatggaatgatggttcgaatactttatgcatacgtccc
ggtgtggttgtaacttatgatcgtaattatgtttctaatcaattgttacgtgagcatggt
atcaaagttattgaaattcctggaagtgaacttgtacgtggtcgaggaggccctcgatgt
atgagtcaaccttaataagagaagatctatag

50   Sequence 2274
MDDKYPFYLDPMPNLYFTRDPQASIGRGMTVNRMFWRARRRESIFISYILKHHPRFKDEN
IPLWVDRDCPFNIEGGDELVLSKDVLAIGISERTSAQAIERLARRIFKDPLSTFKKVVAI
EIPTSRTFMHLDTVCTMIDYDKFTTHSAILKSEGNMNIFIIEYDDKAEDIKIQHSSHLKQ
TLEEVLDVDEITLIPTGNGDIIDGAREQWNDGSNTLCIRPGVVVTYDRNYVSNQLLREHG
55   IKVIEIPGSELVRGRGGPRCMSQPLIREDL*

Sequence 2275
Contig_0734_pos_4456_5061,
is similar to (with p-value 2.0e-70)

>gp:gp|AJ001330|LSAJ1330_2 Lactobacillus sake DNA encoding t
he arginine-deiminase pathway genes. NID: g2764610.
atgaaagggcgcatcccatgaaaaacatcaagaaaccctttgatttaaaaggtaagtca
ttgctaaaagagtatgatcttacaggtgaagaatttgaaggtctaatcgattttgctatg
5   acattaaaaaaatataaacaacaaggcacaccacatcgatatttagagggtaagaatatt
gctttactcttcgaaaagacatctactcggacgcgtgccgcatttacagttgcatctatt
gatctaggtgcacaccctgaatttttagggaaaaatgatattcaattaggaaaaaaagaa
tctgttgaggatactgctaaagtttaggcagaatgtttgatggaattgaatttagaggt
ttttcccaaaaaactgttgaacaattggccgaattctctggagtaccagtatggaatggg
10   ttaactgatgattggcatcctacacaaatgttagctgattatatgacaattaaagaaaat
tttggatatttaaaaggcatcaacctaacttatgtaggaaacggacgtaataatgttgca
cattcgcttatggtggcggtgctttcaatattcaatctgatatgggaggccacgctggag
gattag

15   Sequence 2276
MKGAHPMKNIKKPFDLKGKSLLKEYDLTGEEFEGLIDFAMTLKKYKQQGTPHRYLEGKNI
ALLFEKTSTRTRAAFTVASIDLGAHPEFLGKNDIQLGKKESVEDTAKVLGRMFDGIEFRG
FSQKTVEQLAEFSGVPVWNGLTDDWHPTQMLADYMTIKENFGYLKGINLTYVGNGRNNVA
HSLMVAVLSIFNLIWEATLED*
20

Sequence 2277
Contig_0734_pos_5331_6341,
is similar to (with p-value 4.0e-96)
>gp:gp|Y17554|BLY17554_3 Bacillus licheniformis arcA, arcB,
25   arcC and arcD genes. NID: g3687415.
gtggcatcaatattattgtggagtgtgcacttttaatattaaaaggagttgagatggct
gctttaatcaatagtattgtcactattacaaaattaattcctatcttattagtgattata
tgtatgattgtagcctttaattttaatacattcaagataggtttctttggaatggatgga
tatggatcattatcatttcatttcgctaatacgatgtcacaagttaaaagtacgatgtta
30   gtgacagtttgggtatttattggtattgaaggtgctgttgtattctccggaagagctaaa
aataaaaaagatgttggaactgccactgttatcggacttattcagtattgctcatttac
ttcttgctgactgtattagcacaaggtattgtaattcaaaatcatatttctaaacttgag
gcaccatcaatggcacaaattttagcttatattgttggtgattggggagcaacatttgtc
aatattggtcttattatttcagtattaggtgcatggctaggttggacattacttgccgga
35   gagttgccttttattgtagctaaagatggtttgtttcctaagtggtttgcaaaagaaat
aaaaatggggcacctatgaacgccttattcattactaatgtgttagttcaaatattcctt
attagtatgctgtttaccaaaagtgcttatcattttgcattttctctcgcggcaagtgct
atactttatccatacatgttcagtgcgttttatcaggtgaaatatacaatagaacacaag
ttaactgcaacgcctaaacaatggattataggaattctagcatctatttatgcaatttgg
40   ctggtatacgcgtcaggtatagattacttgttacttaccatgttgctctatattccgggg
attatcgtctatgttgttgttcagaaaaataatcaaaagcgacttacacaatttgactat
attttcttcagtcttatcgttattttagcattgatagggttattacgatga

Sequence 2278
45   VASILLWSVHFLILKGVETAALINSIVTITKLIPILLVIICMIVAFNFNTFKIGFFGMDG
YGSLSFHFANTMSQVKSTMLVTVWVFIGIEGAVVFSGRAKNKKDVGTATVIGLISVLLIY
FLLTVLAQGIVIQNHISKLEAPSMAQILAYIVGDWGATFVNIGLIISVLGAWLGWTLLAG
ELPFIVAKDGLFPKWFAKENKNGAPMNALFITNVLVQIFLISMLFTKSAYHFAFSLAASA
ILYPYMFSAFYQVKYTIEHKLTATPKQWIIGILASIYAIWLVYASGIDYLLLTMLLYIPG
50   IIVYVVVQKNNQKRLTQFDYIFFSLIVILALIGLLR*

Sequence 2279
Contig_0734_pos_2115_1636,
putative peptide of unknown function
55   gtgccagatcatatagagaaagtagtggtcgtagtaaatcctcaaatgtccaccataaag
agaataattaatcaaactgatattaacacaatccaattacatggaaatgaaagcattcaa
ttaattagaaatattaagaaacttaattcaaaaataagaatcataaaagcaattccagca
acaagaaatttaaataataacattcaaaagtataaagatgagatagacatgtttattata
gatacaccatcaatcacatacggagggacaggtcaaagttttgactggaaattattaaaa

562

aaaataaagggcgttgattttctcattgcgggtggtttggattttgaaaagataaaacga
ttagaaatatattcatttggacaatgtggttatgacatctcaactggcattgagtcacat
aatgaaaaagattttaataagatgactcgaatattaaaattttttgaaaggagacgaatga

5

Sequence 2280
VPDHIEKVVVVVNPQMSTIKRIINQTDINTIQLHGNESIQLIRNIKKLNSKIRIIKAIPA
TRNLNNNIQKYKDEIDMFIIDTPSITYGGTGQSFDWKLLKKIKGVDFLIAGGLDFEKIKR
LEIYSFGQCGYDISTGIESHNEKDFNKMTRILKFLKGDE*

10

Sequence 2281
Contig_0734_pos_1632_424,
is similar to (with p-value 0.0e+00)
>sp:sp|Q01998|TRPB_LACLA TRYPTOPHAN SYNTHASE BETA CHAIN (EC

15   4.2.1.20). >pir:pir|S35129|S35129 tryptophan synthase (EC 4.
2.1.20) beta chain - Lactococcus lactis subsp. lactis >gp:gp
|M57193|LACTRPOP_7 L. lactis trpE, trpG, trpD, trpF, trpC, t
rpB trpA genes, complete cds. NID: g149514.
atgaaaattcaaacagaagtagatgaattgggctttttcggtgaatatggtggccaatat

20   gtacctgaaacattgatgccagctattattgaacttaaaaaagcatatgaggacgcgaaa
tcagatactcacttcaagaaagaatttaattattatttaagtgaatatgttggtagagaa
acgcctttaacatttgctgaatcatacacaaaattgttaggtggtgccaaaatatatctt
aaaagagaagacttaaatcacactggtgctcataaaattaataacgcgataggacaggca
ctattagctaaaaggatgggggaaaactaaattagtagccgaaacaggtgctggtcaacat

25   ggtgtagcaagtgccaccatcgctgctttattcgatatggatcttattgttttcatggga
agtgaagatatcaaacgtcaacaacttaacgtatttagaatggaattgctaggagctaaa
gtagtgtctgtgtcagatgggcaaggaacactatcagatgctgtaaataaagctttgcaa
tattgggtgaatcatgtcgaggatacacattatttattaggctcagcgttgggacctgat
ccgtttccaactatggtcagagattttcagagtgtgattggtaatgaaattaaaagccaa

30   attttaagtaaagaaggacgacttccagatgcgttagtcgcgtgtgttggtggaggatcc
aattcaataggtacgttctatccatttatacaagatgatgttaaattatatggggtagaa
gctgcgggaaaaggaagtcatacgcataatcatgctttagctatagggaaaggtaaacca
ggtgtattacatggttccaaaatgtaccttattcaaaatgatgatggacaaattgaattg
gcacactctatatcagcgggactagattatccaggtattggacctgaacattcgtattat

35   aatgatattggtcgtgtatcatatgtaagtgctacagataatgaagctatggaagcactt
ataacattctcaaaagttgaaggtatcattccagcaattgaaagtgcacatgcattgagt
tatgttgaaaaattagcgccaaatatggatgaaaaagaaattattgttgtgactatttca
ggtcgtggagataaagatatggaaacaattaaacaatacaaagaaacggtggtgaacaa
aatgagtaa

40

Sequence 2282
MKIQTEVDELGFFGEYGGQYVPETLMPAIIELKKAYEDAKSDTHFKKEFNYYLSEYVGRE
TPLTFAESYTKLLGGAKIYLKREDLNHTGAHKINNAIGQALLAKRMGKTKLVAETGAGQH
GVASATIAALFDMDLIVFMGSEDIKRQQLNVFRMELLGAKVVSVSDGQGTLSDAVNKALQ

45   YWVNHVEDTHYLLGSALGPDPFPTMVRDFQSVIGNEIKSQILSKEGRLPDALVACVGGGS
NSIGTFYPFIQDDVKLYGVEAAGKGSHTHNHALAIGKGKPGVLHGSKMYLIQNDDGQIEL
AHSISAGLDYPGIGPEHSYYNDIGRVSYVSATDNEAMEALITFSKVEGIIPAIESAHALS
YVEKLAPNMDEKEIIVVTISGRGDKDMETIKQYKENGGEQNE*

50   Sequence 2283
Contig_0734_pos_0_404,
is similar to (with p-value 8.0e-20)
>sp:sp|P17166|TRPA_LACCA TRYPTOPHAN SYNTHASE ALPHA CHAIN (EC
4.2.1.20). >pir:pir|S42347|JS0344 tryptophan synthase (EC 4

55   .2.1.20) alpha chain - Lactobacillus casei >gp:gp|D00496|LBA
TRP_6 Lactobacillus casei DNA, trp operon (trpD, trpC, trpF,
trpE, trpA), complete cds. NID: g216754.
atgggtgatttaaattttattcatcatttaaaaaacattaactgagaatggagcagacatt
gttgaaattggtgtgccattttctgatcctgttgcagatggacctataatcatgaaagca

gggcgcaacgctattgacgagggttcaaacattaaattcatttttgatgaattaataaaa
aataaaaatactatttcatctaagtatgtattaatgacttattataatattctaagtgct
tatggagaagaattattttggataagtgtgatgaagctggtgtttatggtttaattatt
ccagatttaccttacgaacttacaaaaaagtttaaaaaagattttatcatcattctgtt
5   aaaataatatcgttaattgccatgaccgcaagtgatgctaggat

Sequence 2284
MGDLNFIHHLKTLTENGADIVEIGVPFSDPVADGPIIMKAGRNAIDEGSNIKFIFDELIK
NKNTJSSKYVLMTYYNILSAYGEELFLDKCDEAGVYGLIIPDLPYELTKKFKKDFJHHSV
10   KIISLIAMTASDARX

Sequence 2285
Contig_0735_pos_1041_2267,
is similar to (with p-value 3.0e-89)
15   >gp:gp|AF024571|AF024571_1 Staphylococcus aureus high affini
ty proline permease (putP) gene, complete cds. NID: g2565310

atgagtttagtatatgacaaacttactcaagatcaaccgtatcattcgtggtttaatatt
gttgagcatttcttaccttctgatagtcatgatttgttagatattggttgcggtactggc
20   aacttaacacaattactaacgtcactaggtgaagtcactggtatggatattagtgtagat
atgttatcaatagctagacaaaaaacaaatcaagtgaagtggatcgaaggtaatatgact
cactttaatttgaacaaaaaatttaatatgattacaatattttgtgattcactgaattat
ttagaaacattaaatgacgtaaaaatgacattcgaaagagtgtatcaacatttaaataaa
aatggtgtttttattttgatgtacatactgttcataaaatgaaaacattatttaataat
25   aaaagttatattgatgaatctgataatgtttttgtaggttgggatgcaatatgtggggat
gaaccattagtaagatttatgtctatcaaatcacataaattactaccaaaagccagacgt
ctgggaataagttggatggcagtcggtctattaggagcgattggtgtaggattaacagga
attjtcattatatctgaaagacatattaaatcagaagatcctgaaacactatttatjtgtg
atgagtcaaatattatttcatccgcttgtaggtggatttttattagcagccatccjtgct
30   gcaataatgagtactatctcttcacaattactagtaacatcaagttctttaactgaagat
ttctataaactaatcagaggttcagataaagcatcatcacaccaaaaagagtttgttttg
attggacgcttatcagttctacttgttgcgatagttgctattacgattgcttggcatcca
aacgatacaatactaaatttagttggtaatgcttgggctggtttttggagctgcatttagt
cctttagtactctactctttatattggaaagatttaacacgtgcaggagctattagcgga
35   atggtagctggtgctgtggttgttattgtttggatttcttggataaaacccttggctaca
atcaatgcattctttggtatgtatgaaatcattccaggtttcataattagcgtattgatt
acctacatcgtaagtaaattaacaaaaaaacctgatgattatgttattgaaaatcttaat
aaagttaaacacatcgttaaagaataa

40   Sequence 2286
MSLVYDKLTQDQPYHSWFNIVEHFLPSDSHDLLDIGCGTGNLTQLLTSLGEVTGMDISVD
MLSIARQKTNQVKWIEGNMTHFNLNKKFNMITIFCDSLNYLETLNDVKMTFERVYQHLNK
NGVFIFDVHTVHKMKTLFNNKSYIDESDNVFVGWDAICGDEPLVRFMSIKSHKLLPKARR
LGISWMAVGLLGAIGVGLTGISFISERHIKSEDPETLFIVMSQILFHPLVGGFLLAAILA
45   AIMSTISSQLLVTSSSLTEDFYKLIRGSDKASSHQKEFVLIGRLSVLLVAIVAITIAWHP
NDTILNLVGNAWAGFGAAFSPLVLYSLYWKDLTRAGAISGMVAGAVVVIVWISWIKPLAT
INAFFGMYEIIPGFIISVLITYIVSKLTKKPDDYVIENLNKVKHIVKE*

Sequence 2287
50   Contig_0735_pos_4945_3824,
is similar to (with p-value 0.0e+00)
>gp:gp|AJ011676|BST011676_1 Bacillus stearothermophilus lig
gene. NID: g3688228.
atgggttatacgcaaaaatctccaagatgggcgattgcttataaatttccagctgaagaa
55   gttattacaaaattattggatattgagctaagtattgggcgtacgggtgttgtgacacca
actgcaattctagaacctgtaaaagtagctggtactacagtttcaagagcctcacttcat
aatgaagatttaatacatgaaagagatatacgtatcggagatagtgttgttattaaaaaa
gccgggggacatcatccctgaagttgtaaaaagtattttagatagacgacctaacgaatcg
gaaatttatcatatgccaacacattgtcctagttgtggacatgaattagttcgtattgaa ·

ggagaagttgctttacgttgtattaatccaaaatgtcaggcacagcttattgaaggactt
atacatttcgtttcaagacaagcgatgaatatagatggtttaggtactaaaattattcat
cagctatacgaaaatcagttaatcaaagatgtcgcagatattttctatttgaaagaagaa
gatttattaccattagagcgaatgggaaagaagaaagttgataatcttttattagcgata

5   gaaaaatctaaagaacagtcattagagcatttattatttggacttggtattagacattta
ggtgtaaaagctagtcaagtacttgctgagcgatatgaaacgatggatcaacttttttaaa
gtaactgaaagtgaattaattgaaattcaagatattggagataaacttgcacaatctgtt
gtaacatatctcgaaaatagtgatattcgttcattaattgaaaaattaagtaataaaaat
gttaatatgtcttataaaggaattaaaacaactgaaatcgaaggtcatcctgattttagt

10   gggaaaacaattgtattaacagggaaactcgagcaaatgacgagaaatgaagcatctgaa
tggttgaaaatgcaaggtgctaaagttacaagcagcgtgactaaaagtactgatattgtc
atagctggagcagatgcagggtctaaattagccaaagctgagaagtatggtactgaaatt
tggactgaagcagcatttattgaaaaacaaaatggaatctaa

15   Sequence 2288
     MGYTQKSPRWAIAYKFPAEEVITKLLDIELSIGRTGVVTPTAILEPVKVAGTTVSRASLH
     NEDLIHERDIRIGDSVVIKKAGDIIPEVVKSILDRRPNESEIYHMPTHCPSCGHELVRIE
     GEVALRCINPKCQAQLIEGLIHFVSRQAMNIDGLGTKIIHQLYENQLIKDVADIFYLKEE
     DLLPLERMGKKKVDNLLLAIEKSKEQSLEHLLFGLGIRHLGVKASQVLAERYETMDQLFK

20   VTESELIEIQDIGDKLAQSVVTYLENSDIRSLIEKLSNKNVNMSYKGIKTTEIEGHPDFS
     GKTIVLTGKLEQMTRNEASEWLKMQGAKVTSSVTKSTDIVIAGADAGSKLAKAEKYGTEI
     WTEAAFIEKQNGI*

     Sequence 2289
25   Contig_0735_pos_3414_2596,
     is similar to (with p-value 1.0e-49)
     >gp:gp|Z99107|BSUB0004_111 Bacillus subtilis complete genome
      (section 4 of 21): from 600701 to 813890. NID: g2632866. >g
      p:gp|Y15254|BSYERABCD_7 Bacillus subtilis 13kB DNA fragment,

30   from yerA to sapB gene. NID: g2577959. atgagcgaaaaagaaaagaaaagcaaaaatgctaatgagaatcttggactcaatccatct
     cacaatggtgaaacagatgaagagaaaatagctaaaaattctccagcctatctttcaaat
     atactcgagcaggatttttatggaaatagtgattctaaaggtaaaaaatataaaagggatg
     acaattggtttagctatgaatagtgtttattattacaaaaaaagagaaagatggcgaaaca

35   tttagtaaagatttatctgataaagagattgaaaagcaaggtaaacagatggctagtgaa
     atgctttctcgtttacgtgagaatagtgatttgaaagatattcctattcattttgctatc
     tataaacaatcaagtcaagattccattacaccaggtgaatttatagttggtactacggtt
     gaagagggtaaaactaaaattaactcatgggataatattaatgaaaaagcagccttaatt
     ccttcgtcaactgcagctgattatgatgaaacgttgaataataactttaaacagtttaat

40   gataatttgcaatcgtattttttcaaacttcacacaagcagttggtaaggttaaatcgta
     aataaaaaagctaaacaacttacagttgatttgcctatagattattacggacaggcagaa
     acgataggtattacacaatatgttacagagcaagccgaaaaatattttgataaactagat
     gagtatgaaattagaatcaaagatggaaatactccacgtgctctcattagtaaaactaaa
     gacgataaagaaccacaagttcatatctatcataattag

45
     Sequence 2290
     MSEKEKKSKNANENLGLNPSHNGETDEEKIAKNSPAYLSNILEQDFYGNSDSKGKNIKGM
     TIGLAMNSVYYKKEKDGETFSKDLSDKEIEKQGKQMASEMLSRLRENSDLKDIPIHFAI
     YKQSSQDSITPGEFIVGTTVEEGKTKINSWDNINEKAALIPSSTAADYDETLNNNFKQFN

50   DNLQSYFSNFTQAVGKVKFVNKKAKQLTVDLPIDYYGQAETIGITQYVTEQAEKYFDKLD
     EYEIRIKDGNTPRALISKTKDDKEPQVHIYHN*

     Sequence 2291
     Contig_0735_pos_858_538,
55   putative peptide of unknown function
     atgtttatgtgccacttctttaacagatctagcaatggattgaacttgacgttcattgtt
     cccgtgacagactacaaaataaaggggataaaacctcttgggaatgatataagaatatta
     ggtgcatcaagtgatcatttgatgattgatttaaataaccaagatcattatcaaatcggt
     gataaaacttcaatttagcttgaattatgaagcactgtctcagagcatgtatatgaaaaat

565

ttaactaagttatatagtagtgattcaaaaatagaatcccttgttcagaacttcgatatg
cctatatattcccagtgctaa

Sequence 2292
5   MFMCHFFNRSSNGLNLTFIVPVTDYKIKGIKPLGNDIRILGASSDHLMIDLNNQDHYQIG
    DKLQFSLNYEALSQSMYMKNLTKLYSSDSKIESLVQNFDMPIYSQC*

Sequence 2293
Contig_0736_pos_4884_0,
10 is similar to (with p-value 3.0e-23)
    >gp:gp|U93876|BSU93876_19 Bacillus subtilis aminoglycoside 6
    -adenylyltransferase (aadK) gene, partial cds, and YrdA (yrd
    A), YrdB (yrdB), hypothetical protein YrdC (yrdC), YrdD (yrd
    D), hypothetical cytochrome P450 protein YrdE (yrdE), ribonu
15 clease inhibitor (yrdF), regulatory protein YrdG (yrdG), hyp
    othetical protein YrdH (yrdH), hypothetical protein YrdI (yr
    dI), amino acid transporter (yrdJ), YrdK (yrdK), LysR family
      regulatory protein YrdL (yrdL), YrdN (yrdN), cation transpo
    rt protein YrdO (yrdO), hypothetical protein YrdP (yrdP), Ly
20 sR family transcription regulator YrdQ (yrdQ), hypothetical
    protein YrdR (yrdR) and hypothetical protein YrkA (yrkA) gen
    es, complete cds. NID: g1934641. >gp:gp|Z99117|BSUB0014_140
    Bacillus subtilis complete genome (section 14 of 21): from 2
    599451 to 2812870. NID: g2634966.
25 gtgacaatcttagcgattgatattggagtgaatgtgggaatagcatcagcaattgtaaca
    attgtgattatacttatttctgaagtgattcctaaatcaattgctgcaacatttcctgat
    aaaatttcaaaacttgtgtatcctatcattcatatatgtgttattgtactcaagcccatt
    acaatcttattaaacaagatgacagatggtattaatcatttactatctcgaggccaacct
    gttgaaaaaagattttctaaagaagaaattcgtacattattaaatattgcgggtagagaa
30 ggtgcatttaatgagatagaaaatactcgacttcaaaacgttatggactttgaacaattg
    aaggttaaggatgttgataccacgcctcgtattaatgttgtagctttttcaaaggaagta
    acatatgacgaagcttatgatacagtgatgaataacccatatacaagatatccagtatat
    gatgaaaatatagatgatatcatcggcgtattccactcaaaatatttattagcttggagt
    aaaaataaagaggacgcaattactaattatgcatcaagccctttatttgtaaatgaacat
35 aatagggcagaatgggtattgcgtaaaatgaccgtttcacgaaaacatttagcgattgtt
    ttagatgaatttggaggtacggatgctatcgtatcgcacgaagatttaatagaagagcta
    cttggtatggatattgaggatgaaa

Sequence 2294
40 VTILAIDIGVNVGIASAIVTIVIILISEVIPKSIAATFPDKISKLVYPIIHICVIVLKPI
    TILLNKMTDGINHLLSRGQPVEKRFSKEEIRTLLNIAGREGAFNEIENTRLQNVMDFEQL
    KVKDVDTTPRINVVAFSKEVTYDEAYDTVMNNPYTRYPVYDENIDDIIGVFHSKYLLAWS
    KNKEDAITNYASSPLFVNEHNRAEWVLRKMTVSRKHLAIVLDEFGGTDAIVSHEDLIEEL
    LGMDIEDEX
45

Sequence 2295
Contig_0736_pos_3556_2282,
is similar to (with p-value 0.0e+00)
>sp:sp|P37949|LEPA_BACSU GTP-BINDING PROTEIN LEPA. >gp:gp|D8
50 4432|BACJH642_108 Bacillus subtilis DNA, 283 Kb region conta
    ining skin element. NID: g2627063. >gp:gp|X91655|BSLEPONF_1
    B.subtilis lepA and hemN genes. NID: g1122397. >gp:gp|Z99117
    |BSUB0014_31 Bacillus subtilis complete genome (section 14 o
    f 21): from 2599451 to 2812870. NID: g2634966.
55 atgtcaaagatatattgcattaaagtaataaaaacgttactatatcaaagatacatagaa
    gtgacaggttataaagatgaaagcgagaaggataaaatggataagcaagaacgatacaat
    agaagagaaaatattagaaatttctccattattgctcatatagaccatggtaaatcgaca
    ttagctgatcgaattttagagaatacaaaatcagttgaaactcgagaaatgcaagatcaa
    ttacttgactctatggatttggaaagagaacgaggcatcactattaaactaaatgctgtt

566

cgattaaaatacgaagctaaagatggagaaacttacacatttcatttgatagatacacca
ggacatgtcgactttacatatgaggtttctcgctcattagctgcatgtgaaggtgcaatt
cttgtagttgatgctgcccaaggtatagaagcacaaaccttagcaaacgtttattagca
ttagataacgatttggaactttgccagttgttaataaaatagacttgcctgcagctgag
5   cccgatagagttaagcaagaattagaagatgttataggtatagatcaagaagatgtagta
cttgcaagtgctaagtcaaatataggtattgaagaaattttagagaaaatagttgatgtt
gtaccagcaccggacggtgatccagaagccccacttaaagcacttatctttgattcagaa
tatgatccatacagaggagtaatatcttcaattcgaattattgatggtgttgttaaagct
ggagataggattaaaatgatggctaccggtaaagaatttgaagttacagaagtcggaatc
10 aatacccctaagcaactaccggtagaagaattaacagttggtgatgtgggttatattatc
gcaagtatcaaaaatgttgatgattctagagtaggtgacacaattactttagctgaaaga
cctgctgacaaaccgttacaaggatataaaaagatgaatccaatggtattttgtggtcta
ttccctattgacaataaagactataatgacctaagagaagctttagaaaaattacaactt
aatgacgcatccttagagtttgaaccagagtcttcacaagcacttggttttggatacaga
15 actggatttttaggaatgttacatatggagattattcaagaaagaattgaaagagaattt
ggtattgaactcattgcaacagcgccttcatcccgcagaagtatcaatgcatgtgtgaac
accattcgcttttaa

Sequence 2296
20 MSKIYCIKVIKTLLYQRYIEVTGYKDESEKDKMDKQERYNRRENIRNFSIIAHIDHGKST
LADRILENTKSVETREMQDQLLDSMDLERERGITIKLNAVRLKYEAKDGETYTFHLIDTP
GHVDFTYEVSRSLAACEGAILVVDAAQGIEAQTLANVYLALDNDLELLPVVNKIDLPAAE
PDRVKQELEDVIGIDQEDVVLASAKSNIGIEEILEKIVDVVPAPDGDPEAPLKALIFDSE
YDPYRGVISSIRIIDGVVKAGDRIKMMATGKEFEVTEVGINTPKQLPVEELTVGDVGYII
25 ASIKNVDDSRVGDTITLAERPADKPLQGYKKMNPMVFCGLFPIDNKDYNDLREALEKLQL
NDASLEFEPESSQALGFGYRTGFLGMLHMEIIQERIEREFGIELIATAPSSRRSINACVN
TIRF*

Sequence 2297
30 Contig_0737_pos_2511_2855,
putative peptide of unknown function
atgcgtagtgtaattaagttggaaaatacatttatattcataatcacaatcgccgtttat
gttaaattagaattttctatatggctgtttttacttttactattagttccagatatattt
atgttaggatatgtgattaatagaaaaacagggagttatgtttacaatattggacacacg
35 tatatcacacctataattatcgcgctattatatttatacattgatgaaaggttactatta
cagattgctttaatatggttagctcatattagtatggatagaactttaggtttcggactc
aaatattcatcagatactgataaaacgataatacaaaagatgtaa

Sequence 2298
40 MRSVIKLENTFIFIITIAVYVKLEFSIWLFLLLLLVPDIFMLGYVINRKTGSYVYNIGHT
YITPIIIALLYLYIDERLLLQIALIWLAHISMDRTLGFGLKYSSDTDKTIIQKM*

Sequence 2299
Contig_0737_pos_6151_4991,
45 putative peptide of unknown function
gtgttttacgacacagacagtactgaagcgatgaaaagtcatatgagtgatttagtatta
ggcaagcaagaacaaattgcttatatcaatcagttagaacgtggacttgaagaaaataaa
attgaaagaaactctaattctaatgagattaatcaagttgagaatgagcttgttcctgac
gaaaccctttgaaaagaaaaggaatatacacaacaagtttagaattacatgaaaaagag
50 aacttgtatgaaaagttaaaagaaacttttgaagaagaacaaacacaaaaaaataaaaga
caaaagttttgagaataggatttattgttttgactattctatcagcagcactttctatt
ttttctttttttcactgcaaatcttattttggtataatatttgctctattaactgtgatt
tttgtagtaggtatcatttttttctagatctaaagcagtagattatagcacagcaataagt
caggaaattaatgatttagaaaaccaactcacgcaacttgaaaaagaatataatcttgac
55 ttcgatttagaatatcaacaacaagttcgtgaacaatggcgtcatgctaaaaaaaataaa
aaaatacttgaagaaaaacatcaatatatcaatcaatcattaacgactgcaaatgagcga
ttagatagtttaaaacatagcatttttaaactactaatagctggcataatagcaggactt
ttatctggtgtagtcaaactaggttgggaggtaatgttcccaccccgaacacctagtaga
gatgccactaacccacctcaacaactcttacaattactaggaataccgtcaaatattacc

```
catctcacatacaattttttctgagcatgcattaccttggataagttttatcgtacactat
agtttttctatcgctattgcaataatctatatttatatcgcaaagaaatatacaaaaatc
acactaggttatggtgctttatttggtatagttatttggattgttttttcatttaatctta
atgccaattatgcatgtcgtaccgaatgcttttgatcaaccatttcagaacacctatca
```
5   `gaattttttggacacattgtttggatgatggttatagaaatggtcagaaggtatttctat`
```
aatattcaattaaataaataa
```

Sequence 2300
VFYDTDSTEAMKSHMSDLVLGKQEQIAYINQLERGLEENKIERNSNSNEINQVENELVPD
10   ETFEKKKEYTQQVLELHEKENLYEKLKETFEEEQTQKNKRQKFLRIGFIVLTILSAALSI
FSFFTANLIFGIIFALLTVIFVVGIIFSRSKAVDYSTAISQEINDLENQLTQLEKEYNLD
FDLEYQQQVREQWRHAKKNKKILEEKHQYINQSLTTANERLDSLKHSIFKLLIAGIIAGL
LSGVVKLGWEVMFPPRTPSRDATNPPQQLLQLLGIPSNITHLTYNFSEHALPWISFIVHY
SFSIATATIYIYIAKKYTKITLGYGALFGIVIWIVFHLILMPIMHVVPNAFDQPFSEHLS
15   EFFGHIVWMMVIEMVRRYFYNIQLNK*

Sequence 2301
Contig_0737_pos_3491_3075,
putative peptide of unknown function
20   atggaattatatagtatgcctatgtttaataaattttttagttaacgacatagataaatca
```
tcggaatggtatcaagagaatttaggtttttaaaagtattttttaaatttaaaaatgaacaa
aatcaaattttaatggagcatttacgattagcaaaatatcaagatttgatgttaatttct
ggcaaacagtttgaagtcggtaatgcagtttatacaaatatacttgtaccaaatattcga
atttttaaaacaacgaataccttctcaatatatcgtggaagatcttgaagaaaaaccatgg
```
25   `aattctattgaaatgacaattaaagatttagataatcatttaattacgcttacacaaagt`
```
aacataaaaaatgaagaatttaatgctttgatgcaacatacttcaaaaacattttaa
```

Sequence 2302
MELYSMPMFNKFLVNDIDKSSEWYQENLGFKSIFKFKNEQNQILMEHLRLAKYQDLMLIS
30   GKQFEVGNAVYTNILVPNIRILKQRIPSQYIVEDLEEKPWNSIEMTIKDLDNHLITLTQS
NIKNEEFNALMQHTSKTF*

Sequence 2303
Contig_0738_pos_2962_3540,
35   is similar to (with p-value 4.0e-47)
>sp:sp|P42085|XPT_BACSU XANTHINE PHOSPHORIBOSYLTRANSFERASE (
EC 2.4.2.-). >pir:pir|S51309|S51309 xanthine phosphoribosylt
ransferase - Bacillus subtilis >gp:gp|L77246|BACYACA_2 Bacil
lus subtilis (YAC10-9 clone) DNA region between the serA and
40   kdg loci. NID: g1256615. >gp:gp|Z99115|BSUB0012_148 Bacillu
s subtilis complete genome (section 12 of 21): from 2195541
to 2409220. NID: g2634478. >gp:gp|X83878|BSXPTPBUX_1 B.subti
lis xpt and pbuX genes. NID: g633168.
```
gtggagtcgttaggacgaaaagtcaaagaagatggcgttgtcatcgatgagaaattttg
```
45   `aaggtagatggattttttaaatcatcaaattgatgcaaagttgatgaatgatgtaggtaaa`
```
acattttatgagtctttcaaagacgctggtattactaaaattttaactattgaagcttct
ggtattgcgcctgctattatggcttctttttcattttgatgttccttgtctatttgctaaa
aaagctaaacctagtactttgaaagatggctttttatagcacggatattcattcatttaca
aaaaataaaacgagtacagtcattgtatctgaagaattttttaggtgcagacgataaagta
```
50   `cttatcattgatgactttttagctaatggtgatgcttcgctaggtcttaatgacattgta`
```
aaacaagcaaatgcgacgacagttggcgtgggtattgtggttgaaaaaagtttccaaaat
ggtcgccaacgtttagaagatgcaggcttatatgtatcttcactttgtaaggtagcttca
ttaaaaggcaataaggtaactctttttaggtgaagcgtaa
```

55   Sequence 2304
VESLGRKVKEDGVVIDEKILKVDGFLNHQIDAKLMNDVGKTFYESFKDAGITKILTIEAS
GIAPAIMASFHFDVPCLFAKKAKPSTLKDGFYSTDIHSFTKNKTSTVIVSEEFLGADDKV
LIIDDFLANGDASLGLNDIVKQANATTVGVGIVVEKSFQNGRQRLEDAGLYVSSLCKVAS
LKGNKVTLLGEA*

Sequence 2305
Contig_0738_pos_3579_4808,
is similar to (with p-value 8.0e-95)
>sp:sp|P42086|PBUX_BACSU XANTHINE PERMEASE. >pir:pir|S51310|
S51310 xanthine permease - Bacillus subtilis >gp:gp|L77246|B
ACYACA_3 Bacillus subtilis (YAC10-9 clone) DNA region betwee
n the serA and kdg loci. NID: g1256615. >gp:gp|Z99115|BSUB00
12_147 Bacillus subtilis complete genome (section 12 of 21):
 from 2195541 to 2409220. NID: g2634478. >gp:gp|X83878|BSXPT
PBUX_2 B.subtilis xpt and pbuX genes. NID: g633168.
atgtatgcaggggctattcttgttcctattattgtggggacaagcttaaaattttcagct
gaagaaattgcttatctagttactgttgatatatttatgtgcggggtagcgacatttctt
caagcaaataaagtcacagggactggattaccgattgtactaggatgtacgttactgcc
gttgcacctatgactcatcggtcaaacgaaaggacttgatgtttttatatggttcgctt
ttaatatccggtatcttagttgtttttaattgcacctttttctcttatttagttaaattc
tttccacctgttgtaacaggaagtgttgtgacaattattggaatcaatttaatgccagtt
gcaatgaattacttggcaggtggtgaaggagcgaaaaactatggcgatactaagaattta
atattaggtggtgttacactactcattattcttattttgcaaagatttacaaagggcttc
ttgaaatcaattgcgatacttataggattagcaataggtactgctttagctggtatattt
ggaatggttgatatcaaacaagtgggtgatgcacattggtttggtttccctgtgccattc
agattttctggcttcggatttgatgtcagctcaatacttgtattttttcattgttgcagtt
gtaagtttaattgaatctactggtgtctatcatgcactgagtgaaattactggtagaaaa
ctagaaagaaaagattttcgaaaagggtacactgcggaggtctagcaatcattttaggt
tcaatattataatgcgttcccttacactgcatattcccaaaatgtaggtcttgtttcttta
tcaggagctaaaaagaacaatgtgatatatggaatggttattcttttactaatttgcggt
tgtatacctaaattaggtgctttagctaatattattccattgccggttttaggtggagca
atgatagcaatgtttggaatggttatggcatacggcgttagtattttgggtaacattaat
ttccaaaatcaaaataatttattaattattgcaatttcagtagggttaggtgctggtatt
agtgcagtacctcaagcatttaaaggattaggagaacaatttgcttggttaactcaaaat
ggtatagtgcttggcgcaatttctgcaatcatcttaaatttctttttttaatggtataaag
tataaacaaactgaagaaaatgtgaaataa

Sequence 2306
MYAGAILVPIIVGTSLKFSAEEIAYLVTVDIFMCGVATFLQANKVTGTGLPIVLGCTFTA
VAPMILIGQTKGLDVLYGSLLISGILVVLIAPFFSYLVKFFPPVVTGSVVTIIGINLMPV
AMNYLAGGEGAKNYGDTKNLILGGVTLLIILILQRFTKGFLKSIAILIGLAIGTALAGIF
GMVDIKQVGDAHWFGFPVPFRFSGFGFDVSSILVFFIVAVVSLIESTGVYHALSEITGRK
LERKDFRKGYTAEGLAIILGSIFNAFPYTAYSQNVGLVSLSGAKKNNVIYGMVILLLICG
CIPKLGALANIIPLPVLGGAMIAMFGMVMAYGVSILGNINFQNQNNLLIIAISVGLGAGI
SAVPQAFKGLGEQFAWLTQNGIVLGAISAIILNFFFNGIKYKQTEENVK*

Sequence 2307
Contig_0738_pos_4846_6312,
is similar to (with p-value 0.0e+00)
>sp:sp|P21879|IMDH_BACSU INOSINE-5'-MONOPHOSPHATE DEHYDROGEN
ASE (EC 1.1.1.205) (IMP DEHYDROGENASE) (IMPDH) (IMPD). >pir:
pir|S12623|DEBSMP IMP dehydrogenase (EC 1.1.1.205) - Bacillu
s subtilis >gp:gp|X55669|BSIMPDE_1 Bacillus subtilis guaB ge
ne for IMP dehydrogenase. NID: g39958.
atgtgggaaaataaatttgctaaagaatctttaacattcgacgacgtgttactcattcca
gctgcatcagatgtttttaccaagcgatgttgacttaagtgtcaaattatcagataagatc
aagttaaacattcctgttatctcagcaggtatggatacagtaactgaatcaaaaatggca
attgctatggctcgacaaggcggtttaggtgttattcataagaatatgggcgtcgaagag
caagctgatgaggtacaaaaggttaaacgttcagaaatggtgttatttctaacccgttc
ttcttaacaccggaagaaagtgtgtatgaggctgaagcattaatgggtaaataccgtatc
tctggtgtacccattgtcgataatcaagaggatcgcaagttgattgggattttaacaaat
cgtgatttacgtttttattgaagattttttcaattaaaatatcagatgtaatgacgaaagat
aatttaataacagctccagttggtacgactttagatgaagccgaggctattcttcaaaaa

```
cataagattgagaaacttccattagtagaaaatggtcgtttagaaggattaatcactatt
aaagatattgaaaaagtacttgaattcccatatgcagctaaagatgaacatggcagattg
ttagctgcggcagcaatcggtacgtctaaagatactgaaattcgtgcacaaaaactagtt
gaagctggcgtagatgcattaattattgatacagctcatggtcattctaaaggcgttatt
```
5   `aatcaagttaaacacatcaaggaaacatatcctgaaattactgttgtcgctggtaacgta`
```
gcgactgcagaggcaacacgtgctttatttgaagcgggtgccgatgttgttaaagtaggt
attggtccaggctcaatttgcacaacacgtgttgttgcaggtgtaggtgtgcctcaaatt
acagcagtttatgattgtgctacagaagcccgtaagcatggtaaggctattattgctgat
ggtggtattaagttctcaggtgatattatcaaagcattagctgctggtggtcatgcggtt
```
10 `atgttaggtagtttgttagctggtacagaagaaagtcctggtgcaactgaagtattccaa`
```
ggtagacaatataaagtttatcgcggcatgggatctttaggtgctatggaaaaaggttca
aatgatcgttacttccaagaagataaaacaccaagaaaatttgttcctgaaggtattgaa
ggtcgtacagcttataaaggaccattacaagatacaatttatcaacttatgggtggcgtt
agagctggcatgggttatactggttcagaaaacctaaaaaaattacgtgaagaagcacaa
```
15 `tttacacgtatgggaccagctggcttagctgaaagtcatcctcataatgttcaaattacg`
```
aaagaatcaccaaactattctttctag
```

Sequence 2308
MWENKFAKESLTFDDVLLIPAASDVLPSDVDLSVKLSDKIKLNIPVISAGMDTVTESKMA
20 IAMARQGGLGVIHKNMGVEEQADEVQKVKRSENGVISNPFFLTPEESVYEAEALMGKYRI
SGVPIVDNQEDRKLIGILTNRDLRFIEDFSIKISDVMTKDNLITAPVGTTLDEAEAILQK
HKIEKLPLVENGRLEGLITIKDIEKVLEFPYAAKDEHGRLLAAAAIGTSKDTEIRAQKLV
EAGVDALIIDTAHGHSKGVINQVKHIKETYPEITVVAGNVATAEATRALFEAGADVVKVG
IGPGSICTTRVVAGVGVPQITAVYDCATEARKHGKAIIADGGIKFSGDIIKALAAGGHAV
25 MLGSLLAGTEESPGATEVFQGRQYKVYRGMGSLGAMEKGSNDRYFQEDKTPRKFVPEGIE
GRTAYKGPLQDTIYQLMGGVRAGMGYTGSENLKKLREEAQFTRMGPAGLAESHPHNVQIT
KESPNYSF*

Sequence 2309
30 Contig_0738_pos_6473_0,
is similar to (with p-value 5.0e-94)
>gp:gp|U51115|BSU51115_8 Bacillus subtilis CotA (cotA), GabP
(gabP), YeaB (yeaB), YeaC (yeaC), YebA (yebA), GMP syntheta
se (guaA) genes, complete cds, and AIR carboxylase I (purE)
35 gene, partial cds. NID: g2239287. >gp:gp|Z99107|BSUB0004_83
Bacillus subtilis complete genome (section 4 of 21): from 60
0701 to 813890. NID: g2632866.
```
atgactatggaaatggcgaaagagcaagagctgattcttgttttagactttggtagccaa
tataaccagttaattacgcgtcgtatccgtgagatgggcgtttatagtgaattacatgat
```
40 `cacgaaatttctattgaagaaattaaacgaatgaatcctaaaggtatcattctttcaggt`
```
ggtccaaattcagtgtatgaagaggggtcatttaccatcgaccctgaaatttacaattta
ggtattccagttttaggtatatgttatggtatgcaattaacgactaagcttttaggtggt
aaagttgagcgtgccaatgagcgtgaatatggcaaagctacaattaacgctaaatcagat
gaacttttctttggcttaccttctgaacaaacagtatggatgagtcattctgataaagta
```
45 `attgaaattcctgaaggatttgaagtgattgcagatagtccaagtactaattatgcagct`
```
attgaagataaaaaacgtcgcatttacggtgtacaattccatccagaagtacgtcacact
gaatatgggaacgacttactaagaaacttcgtccgccgtgtttgtaattgtacaggtgaa
tggacgatggagaatttcattgaaattgaaattgagaaatccgtcaacaagtaggtaat
cgtaaagtattatgtgcaatgagtggtggagtagattcatccgtagttgctgtactttta
```
50 `cataaggcaatcggcgaccaattaacatg`

Sequence 2310
MTMEMAKEQELILVLDFGSQYNQLITRRIREMGVYSELHDHEISIEEIKRMNPKGIILSG
GPNSVYEEGSFTIDPEIYNLGIPVLGICYGMQLTTKLLGGKVERANEREYGKATINAKSD
55 ELFFGLPSEQTVWMSHSDKVIEIPEGFEVIADSPSTNYAAIEDKKRRIYGVQFHPEVRHT
EYGNDLIPNFVRRVCNCTGEWTMENFIEIEIEKIRQQVGNRKVLCAMSGGVDSSVVAVLL
HKAIGDCLTX

Sequence 2311
```

Contig_0738_pos_2368_1961,
putative peptide of unknown function
atggcaaaaattactgtagtgaataaccaagatgaattatataaagtcatcaatcaaaaa
aaatctgaaggttatttagagacagaattagctgttatcagtaaaagtaagttgcactta
gatgatttacacaactctcaaatctcgttaatggctacaagtggctcatttagtgaccgt
atgtctcgtttacttacaggtgaagatggagaagaaacagtattatctcgttatgattta
actgacaatgaactagaaggatataaacaagatattttaaacgataaaatgctcgttgtt
gc aacagtgaccgttcttctcatgatgaagttgaagataataatgctgcatata gaa
gtggat rtactcattatgccgcagagtctgaagggcctaaagcataa

Sequence 2312
MAKITVVNNQDELYKVINQKKSEGYLETELAVISKSKLHLDDLHNSQISLMATSGSFSDR
MSRLLTGEDGEETVLSRYDLTDNELEGYKQDILNDKMLVVANSDRSSHDEVEDNNAAYKE
VDITHYAAESEGPKA*

Sequence 2313
Contig_0738_pos_1833_1282,
putative peptide of unknown function
atggaatttaaagtgatagaatcagctaaagatccattatttaacgaggcactcaaatta
tatgatgataaattggatattggtttagatgaagatagtaaaattttaaacgctcactt
gaaaataataaaacagaaatgattacgcctttatcgttggaattgaaaaccagactgta
gttagcttagcaactgcacattacgaagcaacaaccaattctgcattttaatttactta
atcgcaaaagaaagcccccaatcatgatgaaagaatgtccttaactttagaggcaatagaa
aaacaattaaaccttttatcacaagaagttcataatagagatattaatttcatcatgtta
gaagttccaaaagaaccctcgactgctaacatcgatgacaagctcagaaatgcactagaa
catcgtcgtcaattcctctttgaaaatcaatttgaaaagcaggacgacattgactatatt
cat ccaaaccaaaccaaaggaaacgccacaaaaagtagatttatttattaaagc aac
atlg at atag

Sequence 2314
MEFKVIESAKDPLFNEALKLYDDKLDIGLDEDSKIFKRSLENNKTENDYAFIVGIENQTV
VSLATAHYEATTNSAFLIYLIAKESPNHDERMSLTLEAIEKQLNLLSQEVHNRDINFIML
EVPKEPSTANIDDKLRNALEHRRQFLFENQFEKQDDIDYIHPNQNQKETPQKVDLFIKAN
IEL*

Sequence 2315
Contig_0741_pos_1807_2565,
putative peptide of unknown function
atggatgacttgaaacaaaatcaatcttctaacgaaaaacctaaaggtaataaaataatt
aatattttgatattcatcggaatgattttattaattcaaatacctattggcgtgtcacta
atagctttaccttttttcagtgaaattcagtaagttaacatccatcgcattaagtatgcta
ataactggtacagcactattaatcatatggttagttaggaattattatttgagtcataca
tatgaaagacaatatcaatcaatgaggggaaaagatatctttattaatattggttttctg
gtattatcaatggttttttagtattctaagtagtgtattaatggtcatatttactggcaac
gatactacagcgaatgagaaagaaatcaatgaaagtttagatttacttttacaaaaagac
catttaccacatatttcaattgttgcaactgttgtttaatgatatgtattataggtccg
tat ttagaggaattactcttccgaggaattttttaaagaaacattatttatgaaatat cga
ttttcgc accattcattatatcttctattatttttagttcacaacatttatcaa aaat
atattttcatatgcaatttattttctaatgggttgtgtattataccttgcctataacaga
agacgtaatatcaaagatagtatgatggttcacatgttgaataattctgtttcaacatta
ccggtatttgttggttatttatggctatattttagatag

Sequence 2316
MDDLKQNQSSNEKPKGNKIINILIFIGMILLIQIPIGVSLIALPFSVKFSKLTSIALSML
ITGTALLIIWLVRNYYLSHTYERQYQSMRGKDIFINIGFLVLSMVFSILSSVLMVIFTGN
DTTANEKEINESLDLLLQKDHLPHISIVATVVLMICIIGPYLEELLFRGIFKETLFMKYR
FWLPFIISSIIFSSQHLSTNIFSYAIYFLMGCVLYLAYNRRRNIKDSMMVHMLNNSVSTL
PVFVGYLWLYFR*

Sequence 2317
Contig_0741_pos_3896_4231,
putative peptide of unknown function
gtgtatatcatgtatgagaatatacaaaatattgaagatttttaccgatttataaatgca
5   cacgctttagctgttgttcatattatgagagataattgtacagtatgtcatgcagtatta
ccccaaattcaagatttactaaaggactatccgaaagcacaattaggtgtgattaatcaa
tctaatgttgaagctattgccggagaactttctattttttacagtacctgttgatttaatt
ttttttgaaagggaaagaaatgcatagacaagcacgtttttatcgatatgcaatcgtttgaa
aaacaattgtatataatgcaaaatgccatcgattaa
10

Sequence 2318
VYIMYENIQNIEDFYRFINAHALAVVHIMRDNCTVCHAVLPQIQDLLKDYPKAQLGVINQ
SNVEAIAGELSIFTVPVDLIFLKGKEMHRQARFIDMQSFEKQLYIMQNAID*

15   Sequence 2319
Contig_0741_pos_4399_5262,
is similar to (with p-value 1.0e-31)
>sp:sp|P42978|YPJC_BACSU HYPOTHETICAL 23.6 KD PROTEIN IN QCR
C-DAPB INTERGENIC REGION. >gp:gp|L38424|BACJOJC_1 Bacillus s
20   ubtilis dihydropicolinate reductase (jojE) gene, complete cd
s; poly(A) polymerase (jojI) gene, complete cds; biotin acet
yl-CoA-carboxylase ligase (birA) gene, complete cds; jojC, j
ojD, jojF, jojG, jojH genes, complete cds's. NID: g755600. >
gp:gp|L47709|BACYPIA_9 Bacillus subtilis (clone YAC15-6B) yp
25   iABF genes, qcrABC genes, ypjABCDEFGHI genes, birA gene, pan
BCD genes, dinG gene, ypmB gene, aspB gene, asnS gene, dnaD
gene, nth gene and ypoC gene, complete cds's. NID: g1146223.
>gp:gp|Z99115|BSUB0012_191 Bacillus subtilis complete genom
e (section 12 of 21): from 2195541 to 2409220. NID: g2634478
30   .
atgaaggaggccaaaatgaagaagcaacagaataaagtccatatgattaatatatcttta
tgccttataggaacattacttattgcaatcgcagtaaatagttttgttataccgggcaat
ttaggtgagggtggttctataggcctttcattaattctgaattatactctaggtatttca
cccgcactcagttccttcatcattaacgcaatattaattattgtaggttggaaatttctc
35   agtagaacgacagcgatttacactgctataacaatcaccgcaagttcaattttttcttgat
ttaactcacacattcggattaggtattcatgataattttattaattcaattttttgcaggt
ttaatgttgggaatcggttctgggttagtaattactgctcatagtacattgggtggtaca
tctgtcattgcacgtatcatttcaaaatatagtgagatgaagacgtcacaagcactactt
atattagatgcaatcatcgtgttatcatttatcgttgtttacctataacgaatgtatta
40   tatactattgtgatgctatttattgttgaaaaatcaatgtctttttgttgttgaaggattt
aatcctaaaaaagctgtgacagttatttcaaaatataataaagaaatcagtgctgatata
tatgaaatgactggaagagggggcaaccttattaagtggtaaaggtgcttaccaaaaaagt
gatacagaagttctatatgccgtggtatcccaaaatcaagttggagcaataaaaaagatt
gttaatcaatatgatgaaaatgccttttttagtgattcatgatgtgcgtgatgtcttaggt
45   aatggatttattaatattaaataa

Sequence 2320
MKEAKMKKQQNKVHMINISLCLIGTLLIAIAVNSFVIPGNLGEGGSIGLSLILNYTLGIS
PALSSFIINAILIIVGWKFLSRTTAIYTAITITASSIFLDLTHTFGLGIHDNFINSIFAG
50   LMLGIGSGLVITAHSTLGGTSVIARIISKYSEMKTSQALLILDAIIVLSFIVVLPITNVL
YTIVMLFIVEKSMSFVVEGFNPKKAVTVISKYNKEISADIYEMTGRGATLLSGKGAYQKS
DTEVLYAVVSQNQVGAIKKIVNQYDENAFLVIHDVRDVLGNGFINIK*

Sequence 2321
55   Contig_0741_pos_7460_8320,
is similar to (with p-value 8.0e-77)
>gp:gp|Z92954|BSZ92954_1 B.subtilis yws[A,B,C,D,E,F,G] and g
erBC genes. NID: g1894764.
gtggttaaacaaaaagcaaatgccctagttaatgagtgtatggctgtaaatccagactat

```
caaattacctttcaaaatgatttagtaaaagcaaatatcggtgtaattgttaatgtgatg
gaagaccatatggatgtcttaggaccgacacttaaagatgtagcgcaagcttttactgca
acaattccatataacgggaaattagttgtaatgaaagataactatactagtttctttgca
aaggaagctaaaaagcgtaattcagaactcattgttgtagataaagacgtcataccagaa
```
5.
```
tcatatttacggaagttcgattatttagtatttcctgataatgtagctattgtgttagga
atagcgcaagcagttggtgtagatgaagaaactgcattacaaggtatgttaaatgcacca
gccgatccaggtgctgttagaattaaatatttccatgcaaatcgcacaaaaaatgtattt
gttaatgcattcgctgctaatgaaccgcagtctacaaaagcgattttaaataaagtggaa
tcatataattatccatacgataagaaaataatcattctcaattgtcgttcagatagggtt
```
10
```
gatagaacacaactctttgttgataacttttaggtgaagtcgattacgatgttct'catt
tgtacagaaaaagtacacaaatggtgacacagtttatggaaactatgccagaaaaaca
tatatcaattatgaaggacgagactttgtagagattgaaaaaggtattctacatgaagct
gagaatgcacttgtattttgtgtaggaaacatccacggcccgggtggtagaatagcggaa
ttcatagaagggatagaataa
```
15

Sequence 2322
VVKQKANALVNECMAVNPDYQITFQNDLVKANIGVIVNVMEDHMDVLGPTLKDVAQAFTA
TIPYNGKLVVMKDNYTSFFAKEAKKRNSELIVVDKDVIPESYLRKFDYLVFPDNVAIVLG

20 IAQAVGVDEETALQGMLNAPADPGAVRIKYFHANRTKNVFVNAFAANEPQSTKAILNKVE
SYNYPYDKKIIILNCRSDRVDRTQLFVDNFLGEVDYDVLICTGKSTQMVTQFMETMPEKT
YINYEGRDFVEIEKGILHEAENALVFCVGNIHGPGGRIAEFIEGIE*

Sequence 2323
Contig_0741_pos_8322_8774,
25 is similar to (with p-value 2.0e-43)
>gp:gp|Z99122|BSUB0019_86 Bacillus subtilis complete genome
(section 19 of 21): from 3597091 to 3809700. NID: g2636029.
>gp:gp|Z92954|BSZ92954_2 B.subtilis yws[A,B,C,D,E,F,G] and g
erBC genes. NID: g1894764.
30 atgataggttcagaattatatttctccttattcgtaggtgtcgtactcagtttgat'attt
gctgagaaatttgggattaatccagcagggttagtcgttccaggttatttagctt'gatt
tttgatcaaccgatcatgttgttatcagtattaatcattagttgcttaacttattttatc
gtaagcaacggtattagtaagtgggttattttatatggtagaagaaaattcgctgccatg
atactgacgggaatggtgattaaatttatatttgatctcttgtacccattgaccccattt
35 gaaatggttgaagtttcaggtataggtgttgtcattcctggtattattgcgaatacaatt
caaaaacaaggtgtagtcattacactttctacaacaatgttattaacatgtattacatat
atcatcttattttatatagtttttattaattaa
```

Sequence 2324
40 MIGSELYFSLFVGVVLSLIFAEKFGINPAGLVVPGYLALIFDQPIMLLSVLIISCLTYFI
VSNGISKWVILYGRRKFAAMILTGMVIKFIFDLLYPLTPFEMVEVSGIGVVIPGIIANTI
QKQGVVITLSTTMLLTCITYIILFLYSFIN*

Sequence 2325
45  Contig_0741_pos_8798_9400,
putative peptide of unknown function
```
atgacaaagaaaaaacgtttatcgcctagtgagtggttgcttaaacaatctaaaagacat
aaaaggaaaaatacactttacacggcaattgtacttttagtagcgttagttctactcata
tttgctgttaaatcaatacaagtagaacctgtaaaaagtgatacgagagacaaagat'agc
```
50  attcgtat'cacctatttaggtaacgtcactttaaataaacatattcgacaaactaacttg
```
aatgatgttttttaaaggtattcaagatactttagatcatagtgattttttcaacaguttca
ttaatagtaaatgattttttcaagaaatcaaaaagataacataaataaaaatattgaaaat
atcatgtttctacgcaagcataatgttaaaagtgttaacttaatcaacgaatctatggat
aatattcaagcgacagcaatgatgagaaaaatagattcccaagcaggttataattttttta
```
55  acaggtaatggttcaaatccaattaatagtaaaactgtacaacaagacattaaaggtaaa
```
aaaatagctaaaacataccatattgttgcagactatctaattgatgtaaatccggattgt
taa
```

Sequence 2326

MTKKKRLSPSEWLLKQSKRHKRKNTLYTAIVLLVALVLLIFAVKSIQVEPVKSDTRDKDS
IRTYLGHVTLNKHIRQTNLNDVFKGIQDTLDHSDFSTGSLIVNDFSRNQKDNINKNIEN
IMFLRKHNVKSVNLINESMDNIQATAMMRKIDSQAGYNFLTGNGSNPINSKTVQQUIKGK
KIAKTYHIVADYLIDVNPDC*

Sequence 2327
Contig_0741_pos_10400_10837,
putative peptide of unknown function
atgagcattgaatttagacatcaatcatggtttgacaatcagtataaagaacaaacttta
tccttcttaacacaacatcaaatcattcatgcagtggtagatgaacctcaagttaaagag
gggagcgttcctttagtaaataggattactagtgaaattgcttttgtacgttatcatgga
cgtaatcattatggttggactaaaaaagatatgactgatcaagaatggcgagatgtaaga
tatttatatgattatagcgatgatgagttagctgacttggctcgtaaagtcgaaatactt
aatcaaaaggctaagaaagtatatgtaatttttaataataactctggcggtcatgcagct
aataatgctaaaaagtatcaaaatattttagacattgattatgaaggtttagcaccgcaa
caattaaaactattttaa

Sequence 2328
MSIEFRHQSWFDNQYKEQTLSFLTQHQIIHAVVDEPQVKEGSVPLVNRITSEIAFVRYHG
RNHYGWTKKDMTDQEWRDVRYLYDYSDDELADLARKVEILNQKAKKVYVIFNNNSGGHAA
NNAKKYQNILDIDYEGLAPQQLKLF*

Sequence 2329
Contig_0741_pos_11189_0,
putative peptide of unknown function
gtgtttatgatatttgtgtcgatattactgatgattcgtcacaaaatcaaaccttttaaa
atttttgacaaacctaaatatgcgcgtacatatgttgatgctgaagggaaaacataccgt
tatagtgtaccacccctgtttgcttttataacaacgttatttattgggctattaacagga
ctgtttggcataggtggaggtgcattgatgacccctcttatgctcatcgtctttagattt
ccaccacatgttgcagtaggcacaagtatgatgattttcttttcaagtgtgatgagt
tcaatagggcacatctttcaaggacatgtggcttggggctattctatcattctcattatt
tcaagtgttataggtgcacaaattgaggacttcatttctcccacatcaggattgaaatat
gagcgccatcaatataacttatacggccaacttcatagaaaagaatattattatgatgac
tcttctt

Sequence 2330
VFMIFVSILLMIRHKIKPFKIFDKPKYARTYVDAEGKTYRYSVPPLFAFITTLFIGLLTG
LFGIGGGALMTPLMLIVFRFPPHVAVGTSMMMIFFSSVMSSIGHIFQGHVAWGYSIILII
SSVIGAQIEDFISPTSGLKYERHQYNLYGQLHRKEYYYDDSSX

Sequence 2331
Contig_0741_pos_9875_9345,
is similar to (with p-value 1.0e-58)
>sp:sp|P23920|SYM_BACST METHIONYL-TRNA SYNTHETASE (EC 6.1.1.
10) (METHIONINE--TRNA LIGASE) (METRS). >pir:pir|S16682|S1668
2 methionine--tRNA ligase (EC 6.1.1.10) – Bacillus stearothe
rmophilus >gp:gp|X57925|BSMETSG_1 B.stearothermophilus metS
gene for methionyl-tRNA synthetase. NID: g39988.
gtggtaagagattacttaatgcgtgaattaccgtttggttctgatggcgtatttacaccg
gaagcctttgttgaaagaacaaattacgatcttgcgaatgatttaggtaatctagtgaat
cgtactatctctatgataaacaaatatttccacggcgaattacctgcataccaaggtcca
aaacatgaattggatgaaaaaatggaagcgatggcgcttgaaactgttaaatcattcaat
gataatatggaaagtttacaattttctgttgctttatcaacagtatggaaatttattagt
cgtacaaacaaatatattgatgaaactcaaccttgggttcttgcaaaagatgaaaatcaa
cgtgagatgcttggtaatgtaatggcacatcttgtcgagaacattcgtttcgctacaatc
ttattacaaccattcttgacgcatgcacctagagagatatttaagcaacttaatattaac
aatccggatttacatcaattagatagtctgcaacaatatggtatgtttttag

Sequence 2332

VVRDYLMRELPFGSDGVFTPEAFVERTNYDLANDLGNLVNRTISMINKYFHGELPAYQGP
KHELDEKMEAMALETVKSFNDNMESLQFSVALSTVWKFISRTNKYIDETQPWVLAKDENQ
REMLGNVMAHLVENIRFATILLQPFLTHAPREIFKQLNINNPDLHQLDSLQQYGMF*

5    Sequence 2333
     Contig_0741_pos_6714_6397,
     putative peptide of unknown function
     atggagaagtcagttaaacttgctgttggcatctatctagcaattattttaattatttgt
     agtatttacctcgcatttatacttattggaagcttaaatggtaaagacatgagtaattct
10   gtttagatactgatcactctcgtatcaacaatacttcaagaaacagtaacgaagatgtt
     acgtcatcaaataatgagtcaaacaatacaaaagcgcactcatttgcaaactctgaatat
     aaagctattaacataaacgaagcatttaaaaataataagcaaattaaaaaagcgaattcg
     agttatcaatactattga

15   Sequence 2334
     MEKSVKLAVGIYLAIILIICSIYLAFILIGSLNGKDMSNSVLDTDHSRINNTSRNSNEDV
     TSSNNESNNTKAHSFANSEYKAININEAFKNNKQIKKANSSYQYY*

     Sequence 2335
20   Contig_0741_pos_2363_2061,
     putative peptide of unknown function
     atgaatggtagccaaaatcgatatttcataaataatgtttctttaaaaattcctcggaag
     agtaattcctctaaatacggacctataatacatatcattaaaacaacagttgcaacaatt
     gaaatatgtggtaaatggtctttttgtaaaagtaaatctaaactttcattgatttctttc
25   tcattcgctgtagtatcgttgccagtaaatatgaccattaatacactacttagaatacta
     aaaaccattgataataccagaaaaccaatattaataaagatatctttccctcattgat
     tga

     Sequence 2336
30   MNGSQNRYFINNVSLKIPRKSNSSKYGPIIHIIKTTVATIEICGKWSFCKSKSKLSLISF
     SFAVVSLPVNMTINTLLRILKTIDNTRKPILIKISFPLID*

     Sequence 2337
     Contig_0743_pos_1628_1969,
35   putative peptide of unknown function
     atgttcacaatcaaacctgtagaaatagctacaactaaaatggcgataacaaatgagact
     acaaatttttttcttgaatagtttcgataataatacaacttctggaatacttgcaccagca
     ccaccaatgatcaaagcaacaactgtcccaagcgacattcctttgaagctaatgcttca
     gctataggtaacattgtttcaggtctaatatacattggaatgcctataacagatgcaatg
40   aatacagatataacgccatcaccacttgcgtattttgtaataaatgtttcgggtacaaag
     ccatatatgaacgctccaataaacacaccaataaataggtaa

     Sequence 2338
     MFTIKPVEIATTKMAITNETTNFFLNSFDNNTTSGILAPAPPMIKATTVPSDIPFEANAS
45   AIGNIVSGLIYIGMPITDAMNTDITPSPLAYFVINVSGTKPYMNAPINTPINR*

     Sequence 2339
     Contig_0743_pos_3389_3778,
     is similar to (with p-value 2.0e-16)
50   >gp:gp|AF016485|AF016485_124 Halobacterium sp. NRC-1 plasmid
       pNRC100, complete plasmid sequence. NID: g2822278.
     atgttaatctataatagcaaattaattaatgaggtgagtataatgttaaatattgaaata
     tacgaagaagctatgtgctgttctactggagtatgtggtccagaaccggatgaaaacta
     ataaaagcgaaccaaatcaatgaatatttaaagcaaatcaaatagaagttcaacgttat
55   aatatgaataataatccgaatgaattcattaaaaatcaagaagttattcgtttaattcaa
     gaaaaaggtgatgaagtttaccaatcacttttatagaaggcggtatagctaaaacgggc
     gcttatattacccaagaagaagccgatgaaattattacagttaatcaaatgagaaatgga
     ggatgctgtggtggagatggatgctgttaa

Sequence 2340
MLIYNSKLINEVSIMLNIEIYEEAMCCSTGVCGPEPDETLIKANQINEYLKQNQIEVQRY
NMNNNPNEFIKNQEVIRLIQEKGDEVLPITFIEGGIAKTGAYITQEEADEIITVNQMRNG
GCCGGDGCC*

5

Sequence 2341
Contig_0743_pos_4248_4676,
putative peptide of unknown function
atgcttgaattaccttctgcatggacagattatttaaatacaacgagtaatgacgcttct
10    tgcttaggtcaattatcaggtttaaatgaaaatagagttaaatataattcagcacttgaa
aaactacgtaaccaagatgatacgaccatgatgttagttgcgagacctactcactcttct
atatatgaaattcaaagagcgcaacaagaattcatcatccactggtggtgtgactacttc
gcctgtatcaggattttttaactcctggtttacctggaacgtcctcttggctacctttcgg
tgcatttggatcaaattcatccttatggcctggcttgatttcttcgccaccataatgaac
15    gatttcatctactggttgttttgttacttttctgttggttcaccttcgccaacttttc
ccctgttaa

Sequence 2342
MLELPSAWTDYLNTTSNDASCLGQLSGLNENRVKYNSALEKLRNQDDTTMMLVARPTHSS
20    IYEIQRAQQEFIIHWWCDYFACIRIFNSWFTWNVLLATFRCIWIKFILMAWLDFFATIMN
DFIYWLFCYFFCWFTFANFFPC*

Sequence 2343
Contig_0743_pos_5242_4406,
25    putative peptide of unknown function
gtggatgatgtgacaaaatatggtccagttgatggagatccgattacgtcaacggaagaa
attccgtttgataaaaaacgcgaatttgatccaaacttagcgccaggtacagagaaagtc
gttcaaaaaggtgaaccaggaacaaaaacaattacaacaccaacaactaagaacccatta
acaggggaaaaagttggcgaaggtgaaccaacagaaaaagtaacaaaacaaccagtggat
30    gaaatcgttcattatggtggcgaagaaatcaagccaggccataaggatgaatttgatcca
aatgcaccgaaaggtagccaagaggacgttccaggtaaaccaggagttaaaaaccctgat
acaggcgaagtagtcacaccaccagtggatgatgtgacaaaatatggtccagttgatgga
gatccgatcacgtcaacggaagaaattccgtttgataaaaaacgcgaatttgatccaaac
ttagcgccaggtacagagaaagtcgttcaaaaaggtgaaccaggaacaaaaacaattaca
35    acaccaacaactaagaacccattaacaggggaaaaagttggcgaaggtgaaccaacagaa
aaagtaacaaaacaaccagtagatgaaatcgttcattatggtggcgaagaaatcaagcca
ggccataaggatgaatttgatccaaatgcaccgaaaggtagccaagaggacgttccaggt
aaaccaggagttaaaaatcctgatacaggcgaagtagtcacaccaccagtggatgatgaa
ttcttgttgcgctctttgaatttcatatatagaagagtgagtaggtctcgcaactaa
40

Sequence 2344
VDDVTKYGPVDGDPITSTEEIPFDKKREFDPNLAPGTEKVVQKGEPGTKTITTPTTKNPL
TGEKVGEGEPTEKVTKQPVDEIVHYGGEEIKPGHKDEFDPNAPKGSQEDVPGKPGVKNPD
TGEVVTPPVDDVTKYGPVDGDPITSTEEIPFDKKREFDPNLAPGTEKVVQKGEPGTKTIT
45    TPTTKNPLTGEKVGEGEPTEKVTKQPVDEIVHYGGEEIKPGHKDEFDPNAPKGSQEDVPG
KPGVKNPDTGEVVTPPVDDEFLLRSLNFIYRRVSRSRN*

Sequence 2345
Contig_0743_pos_2507_1623,
50    putative peptide of unknown function
atgttagattcaattatggaatttataaaaacatttgtgatgttgttttttgaattatta
atactattcatcatcgtgagctttattgtaagtataatccagcagatagtttcagaagaa
aaaataaaacacttttttaagtaaacctaatcaagctattaattatattttgggaatggct
tttggtgcgatgacaccattttgttcttgttctacaattcctatacttgcaggtttatta
55    aattctaaagttccattggccctgcaatgagttttttaattgcgtcacctttaatgaat
ccattaatgatatttatgttatgggccttattaggttggaaagttgctgttgtttacttt
attttactagcactctttagtgtcttaacaggtctagtattttcaaaaatgaatttagct
gaaacttataaaggagtaaatgttaaaggcgatggattttttgctaataaaatgggatct
cgttttaaacaagcattaaatgatgcgtgggcattttttatatccaatgcttccttaccta

tttattggtgtgtttattggagcgttcatatatggctttgtacccgaaacatttattaca
aaatacgcaagtggtgatggcgttatatctgtattcattgcatctgttataggcattcca
atgtatattagacctgaaacaatgttacctatagctgaagcattagcttcaaaaggaatg
tcgcttgggacagttgttgctttgatcattggtggtgctggtgcaagtattccagaagtt
5    gtattattatcgaaactattcaagaaaaaatttgtagtctcatttgttatcgccatttta
gttgtagctatttctacaggtttgattgtgaacatcgttatttaa


Sequence 2346
MLDSIMEFIKTFVMLFFELLILFIIVSFIVSIIQQIVSEEKIKHFLSKPNQAINYILGMA
10   FGAMTPFCSCSTIPILAGLLNSKVPFGPAMSFLIASPLMNPLMIFMLWALLGWKVAVVYF
ILLALFSVLTGLVFSKMNLAETYKGVNVKGDGFFANKMGSRFKQALNDAWAFLYPMLPYL
FIGVFIGAFIYGFVPETFITKYASGDGVISVFIASVIGIPMYIRPETMLPIAEALASKGM
SLGTVVALIIGGAGASIPEVVLLSKLFKKKFVVSFVIAILVVAISTGLIVNIVI*


15   Sequence 2347
Contig_0744_pos_3530_4705,
is similar to (with p-value 2.0e-71)
>sp:sp|P39643|AAT2_BACSU PROBABLE ASPARTATE AMINOTRANSFERASE
(EC 2.6.1.1) (TRANSAMINASE A) (ASPAT). >pir:pir|S39740|S397
20   40 hypothetical protein - Bacillus subtilis >gp:gp|X73124|BS
GENR_86 B.subtilis genomic region (325 to 333). NID: g413923
. >gp:gp|Z99123|BSUB0020_65 Bacillus subtilis complete genom
e (section 20 of 21): from 3798401 to 4010550. NID: g2636240


25   atgaggaggcatagcatggaaatgtcagaaaggctagcttcaattcctgatagctacttt
ggcaaaacaatgggccgtatagttgaacatggtcctttaccacttataaatatggcagtt
ggaattccagatggagaaacgccaaagggtattatcaatcattttttcagaggcgctatgt
attccagaaaatcaaaagtatggtccatttcacggcaaagatgcctttaaacaagctatt
gttaacttctaccaaagacattacgatgttgaattagacaaagaagatgaagtttgtatt
30   ttatatgggactaaaaatggtcttgttgcattacctacttgtgttgttaatcctggtgaa
attgtacttttacctgatccgggatatacagattatttagcggggggtcatgttagctgat
gctaagccactcccttttaaaattgtcgccaccaaattatttgccgaattggaatactata
agtgctaaagttcttgagaagactaagctaatttatttaacatatcccaataatcctacc
ggttcgacagcgacacaagatgattttgatgaagcgattcatcgtttaaaggtactcaa
35   acaaagatagttcatgactttgcatatagtgcttttggatttgacgccaaaaatccaagc
atattagcttctaaaaatgcaaaagatgttgctatcgagatattctctttatctaaaggt
tataatatgtcaggctttcgtgttgggtttgctgttggtaataaaaaaatgattcaagcg
ttaaagaagtatcaaactcatacaaatgcaggtatgtttggagcacttcaagatgctgct
acgtatgcactcaatcattatgatgagtttttagaaaagcaaaatgaaatatttagacgt
40   agacgtgataattttgaatcacaactaaaacatgcacatttaccgtttgttcactctaag
ggaggtatttacatttggttacatacaccccgggttatgatagtgaagcattcgaacag
ttgttattaaaagaaaagtcaatttttagttgcacctggtaaaccatttggtgaaaatggt
aatcaatatgtgagggtttcattggcgctcgatgataaacaattagaagaagcggcgaat
cgcttaacacaattacggtatttgtatgaaagataa
45

Sequence 2348
MRRHSMEMSERLASIPDSYFGKTMGRIVEHGPLPLINMAVGIPDGETPKGIINHFSEALC
IPENQKYGPFHGKDAFKQAIVNFYQRHYDVELDKEDEVCILYGTKNGLVALPTCVVNPGE
IVLLPDPGYTDYLAGVMLADAKPLPLKLSPPNYLPNWNTISAKVLEKTKLIYLTYPNNPT
50   GSTATQDDFDEAIHRFKGTQTKIVHDFAYSAFGFDAKNPSILASKNAKDVAIEIFSLSKG
YNMSGFRVGFAVGNKKMIQALKKYQTHTNAGMFGALQDAATYALNHYDEFLEKQNEIFRR
RRDNFESQLKHAHLPFVHSKGGIYIWLHTPPGYDSEAFEQLLLLKEKSILVAPGKPFGENG
NQYVRVSLALDDKQLEEAANRLTQLRYLYER*


55   Sequence 2349
Contig_0744_pos_4718_5716,
is similar to (with p-value 7.0e-84)
>gp:gp|U31175|SAU31175_1 Staphylococcus aureus D-specific D-
2-hydroxyacid dehydrogenase (ddh) gene, complete cds. NID: g

1644432.
atgactaaaattaaattaatgggtgtcagagaagaagatgaacattatattgaaatgtgg
tcacaacaacatgaagtggaagtggatatgtcgaaagaacagttaactgaagacaatgtc
caatctattgaaggatttgatggactatcattgtctcaaacattaccattatcagaaaca
5    atttataataaattaaatcaacttggaattcggcagatcgctcaacgaagtgctggatttt
gatggttataatttagagttagcatctaaatatggtcttattatatctaatgtgccttcc
tattcacctcgaagcattgctgagtttaccgtgactcaagccatcaatattgtacgtcac
tttaɛtcatattcaaagaaaaatgagattgcacgattttaggtgggaagcatcaatttta
tctcaatcaatcaaagatttaaaggtagcggttattggcacgggacatattggtggcatt
10   gttgcacaaatattctcagaaggatatctatgtgacgttgtagcgtatgatccttttcca
agtgaacatgtgaaaccttacgttacctataaacaaagtataaatgaggcaattaaagag
gcagatattgtcacaatacatatgccgtcaacacaatataacaattacctgtttaatgaa
aacatgtttcaaatgtttaaaaagggtgctgtgtttgtaaattgtgctagaggatcctta
gtagataccaaggctttgttatctgcaatagagcaaggtcaaattaaaggtgcagcactt
15   gatacttatgaatatgaaattggagtatatacgacagatagaagtgaagaaggttttgaat
gacccacttttagaggaattaattactagagaagatattattgttacaccgcatatagca
ttttatactgaagaggcaatcaaacatcttattttttgatgctttagatgcaacaatggaa
gtattaaatactggcacgacggagttaagagtaaattaa


20   Sequence 2350
MTKIKLMGVREEDEHYIEMWSQQHEVEVDMSKEQLTEDNVQSIEGFDGLSLSQTLPLSET
IYNKLNQLGIRQIAQRSAGFDGYNLELASKYGLIISNVPSYSPRSIAEFTVTQAINIVRH
FNHIQRKMRLHDFRWEASILSQSIKDLKVAVIGTGHIGGIVAQIFSEGYLCDVVAYDPFP
SEHVKPYVTYKQSINEAIKEADIVTIHMPSTQYNNYLFNENMFQMFKKGAVFVNCARGSL
25   VDTKALLSAIEQGQIKGAALDTYEYEIGVYTTDRSEEGLNDPLLEELITREDIIVTPHIA
FYTEEAIKHLIFDALDATMEVLNTGTTELRVN*


Sequence 2351
Contig_0744_pos_8456_6108,
30   is similar to (with p-value 0.0e+00)
>sp:sp|P32113|ATKA_ENTFA POTASSIUM/COPPER-TRANSPORTING ATPAS
E A (EC 3.6.1.36). >pir:pir|A45995|A45995 Cu2+-transporting
ATPase (EC 3.6.1.-) - Enterococcus hirae >gp:gp|L13292|ENECO
PPUMP_1 Enterococcus hirae ATPase (copA) gene, complete cds;
35    ATPase (copB) gene, complete cds. NID: g290641.
atgacatgtgctgcgtgctcaaatcgtattgaaaagaaattgaatcgtatgaatcatgtt
caagctaaagtgaatctgactactgaaaaagcaactatcgactatgagtctgacgattat
catctcgaagattttgtagagcaaattcaaagtctcggctatgatgttgcagttgagcaa
gtagaattaaatataaatggtatgacatgtgctgcatgttctaatcgtatagaaaaggtt
40   ctaaatcaaacgcaaggtgtacaacaggcaacagtaaacttaactaccgaacaagcactc
atcaaatattaccctagtgctacgaacacggaagcattaattaagcgtattcaaaatatt
ggatacgatgctgaaactaaaacttcatcaaaagcgcaatcaaatcgtaaaaaacaagag
ttaaaacataaacgcaataaaattaatcatttcagctattttatcgttgccactattatta
gtaatggtggtgcatatctcacctatttccattccatccattttggtcaatccttgggta
45   caattaattctttcaacacctgtccaatttattattgggtggcaattttacgttggcgcg
tataaaaatttgcgaaatggttcagctaacatggatgtattggttgctgttggtaccagt
gccgɔatattttatagcatttatgaaatgatgatgtggctcacacatcaaacacatcac
ccgcatttatattttgaaacaagtgctatttaattacgttaattcttcttggtaaatat
ttagaagcacgtgccaaatcacagactaccaatgcattaagcgaattgttaaatttacaa
50   gcgaaagaagcacgagtaattaaagaaaataaagaaattatgcttccacttgataaagtt
aaagtcggagatactttactaataaaacccggcgaaaagatacctgtagatggcaaagtc
actaaaggtgatacttctattgacgaatccatgctaactggtgagtctatacctgttgaa
aaaagtagtggcgattcagtgattggttctaccatgaataaaaatggttcaatcatgatt
gaagcaactcaagtaggtggtgatactgcattatcacatataattaaagtggttgaggat
55   gctcaaagttctaaagcaccgattcaacgcttagctgatattatttctggatattttgtt
ccgattgtagttagcattgcggttattacttttatcatatggattatattcgttcacccc
gggcaatttgaacctgcacttgtttcagcaatatctgttttagttattgcttgtccctgt
gcacttggtttagcaacgcctacatctattatggtaggtacaggacgtgctgcagaaaat
ggcatattattcaaaggaggccaatttgtagaacgtgcacattatgttgatacaatcgtg


578

```
ctagataaaacaggcacaattactaatggtcaacctgtagtaactgattatgttggtgac
aatgatacattacaacttttagcaagtgctgaaaatgcttcagaacatcctcttgctgat
gctattgttacttatgctaaagataaaggtcttaatttacttgataatgacacttttaaa
tcaattccgggacatggtattaaagctacgattcatcaacaacaaatccttgtgggcaat
cgaaaattaatgaacgattacaatatatctattagtaataaattaaatgaccaattaaat
cactatgaacatttaggtcaaacggcaatgatgattgccgtggataatcaaattaatgga
atcattgctgttgctgatacagtaaaaaatgatgctaaacaagcgataaaagaactaaga
aatatgaatatcgacgtggttatgctgactggtgataacaatcgaacagctcaaaccatc
gccaaacaagttggcattgaacatgtaattgcagaagtgttgcccgaagaaaaggcacat
caaatctctttattacaagacaaaggtaaacaggttgccatggtcggtgatggaattaat
gatgcgcctgcacttgtaaaagccgatattggaatggctataggcactggagctgaggta
gcgattgaagctgcagatattacgattcttggtggtgacttgctattagttccaaaagct
atcaaagcaagtaaagctacgattaaaaatattcgacaaaatttattttgggcatttgga
tataacgtagctggcatcccaatagctgcttgtggtttattagcaccttggattgccggt
gctgctatggcattaagttctgttagcgtagttatgaatgcattaagactgaaaaaaatg
aaactatag
```

Sequence 2352
MTCAACSNRIEKKLNRMNHVQAKVNLTTEKATIDYESDDYHLEDFVEQIQSLGYDVAVEQ
VELNINGMTCAACSNRIEKVLNQTQGVQQATVNLTTEQALIKYYPSATNTEALIKRIQNI
GYDAETKTSSKAQSNRKKQELKHKRNKLIISAILSLPLLLVMVVHISPISIPSILVNPWV
QLILSTPVQFIIGWQFYVGAYKNLRNGSANMDVLVAVGTSAAYFYSIYEMMMWLTHQTHH
PHLYFETSAILITLILLGKYLEARAKSQTTNALSELLNLQAKEARVIKENKEIMLPLDKV
KVGDTLLIKPGEKIPVDGKVTKGDTSIDESMLTGESIPVEKSSGDSVIGSTMNKNGSIMI
EATQVGGDTALSHIIKVVEDAQSSKAPIQRLADIISGYFVPIVVSIAVITFIIWIIFVHP
GQFEPALVSAISVLVIACPCALGLATPTSIMVGTGRAAENGILFKGGQFVERAHYVDTIV
LDKTGTITNGQPVVTDYVGDNDTLQLLASAENASEHPLADAIVTYAKDKGLNLLDNDTFK
SIPGHGIKATIHQQQILVGNRKLMNDYNISISNKLNDQLNHYEHLGQTAMMIAVDNQING
IIAVADTVKNDAKQAIKELRNMNIDVVMLTGDNNRTAQTIAKQVGIEHVIAEVLPEEKAH
QISLLQDKGKQVAMVGDGINDAPALVKADIGMAIGTGAEVAIEAADITILGGDLLLVPKA
IKASKATIKNIRQNLFWAFGYNVAGIPIAACGLLAPWIAGAAMALSSVSVVMNALRLKKM
KL*

Sequence 2353
Contig_0745_pos_341_742,
is similar to (with p-value 5.0e-21)
>sp:sp|P35154|YPUG_BACSU HYPOTHETICAL 29.6 KD PROTEIN IN RIB
T-PACB INTERGENIC REGION (ORFX7). >pir:pir|S45549|S45549 hyp
othetical protein X7 - Bacillus subtilis >gp:gp|L09228|FACDI
A_16 Bacillus subtilis spoVA to serA region. NID: g410114. >
gp:gp|Z99116|BSUB0013_34 Bacillus subtilis complete genome (
section 13 of 21): from 2395261 to 2613730. NID: g2634723.
```
atgtattcatgctataataatttagtgaaatatgcattaacgatgcaatttgatatagag
gtagatgtaatgtatgaagtaaaactcgatgcatttaatggtccattagacttattattg
catctaattcaaaaatatgaaattgatatttatgatatccctatgaaagccttaactgaa
cagtacatgcaatatgttcatgcgatgaatcagctagaaattaatgttgctagtgaatat
ttagttatggcatcagaattactaatgattaaaagtaaattcaacgatgaatataaccaa
aaatccaacactaaaaacatgctgtgcggcttctgcactatgaactgttgtttcaccatt
aactgttgtagtaaaattagcatccatcgctccagtaattga
```

Sequence 2354
MYSCYNNLVKYALTMQFDIEVDVMYEVKLDAFNGPLDLLLHLIQKYEIDIYDIPMKALTE
QYMQYVHAMNQLEINVASEYLVMASELLMIKSKFNDEYNQKSNTKNMLCGFCTMNCCFTI
NCCSKISIHRSSN*

Sequence 2355
Contig_0745_pos_1717_2316,
putative peptide of unknown function
gtgacttctatgtctacgaataatgagattgaatttaaacaaatactagatcaagatact

```
tactcaaaaatctatgaacactatttcaaaaatcaatcaccttttaagcaaactaatttc
tatatcgacacagagaattttaaattaaaacagcatcatgctgctttgcgtataagggta
aaagattatatgtttgaaatgactttaaaagtccctgctgaagttggattgacagaatat
aatcactcagtaaatatagaacctgaacttgatatgtcacttcaactttctcaattaccc
aacgatattagaaatattttagaacaggactttaatattttagaaaatgagcttaaagta
ctaggaaacttaactacctatcgtttagaaaccgattatcaaaatgaattactagtatta
gataagagtgaatatctcggcaaaactgattatgaattagagtttgaagttcattcttat
gatgaaggatattcaaaatttaaaactttacttcaacattttaatcttcaacatcaaaaa
cccttgaataaagtgcaacgttttttttcaagaaaaacaaaatgcaagtgataaagagtaa
```

Sequence 2356
VTFMSTNNEIEFKQILDQDTYSKIYEHYFKNQSPFKQTNFYIDTENFKLKQHHAALRIRV
KDYMFEMTLKVPAEVGLTEYNHSVNIEPELDMSLQLSQLPNDIRNILEQDFNILENELKV
LGNLTTYRLETDYQNELLVLDKSEYLGKTDYELEFEVHSYDEGYSKFKTLLQHFNLQHQK
PLNKVQRFFQEKQNASDKE*

Sequence 2357
Contig_0745_pos_2380_2763,
putative peptide of unknown function
```
atggaacacggtgatattatgtctaaaacaccatatgagttgattggtcaaaaagccttg
tatcaaatgattgatcatttctatcaacttgtcgagaaagattctcgtatcaatcattta
ttcccaggcgatttcaaggaaaccagtcgaaagcagaagcaatttttgacacagtttctt
ggaggtcctgacttatatacccaagaacatggtcatcccatgttaaaacgaagacatatg
gaatttacaattagcgagtatgaacgtgatgcatggcttgagaacatgcatactgctatt
caacacgccaaacttcctgcgggtgtaggcgattacttgtttgagcgattaagacttact
gcaaatcacatggtaaattcctaa
```

Sequence 2358
MEISGIMSKTPYELIGQKALYQMIDHFYQLVEKDSRINHLFPGDFKETSRKQKQFLTQFL
GGPDLYTQEHGHPMLKRRHMEFTISEYERDAWLENMHTAIQHAKLPAGVGDYLFERLRLT
ANHMVNS*

Sequence 2359
Contig_0745_pos_2807_3583,
putative peptide of unknown function
```
atggagaataagagtcgtgaagatactaatctatcacctgttagcaaaatagaaatctat
tcttttttttgatccttttagcaaagattgtttttaaattatctgcaatcttatcaaaatta
agaattgaatataataaatatataaaggtaagacatattttaaacccttctttaaaggta
ttaactaagtgtcaagctcaaagtacttcagattttgacaatattgcacttgcctataaa
gccgctgaacttcaaggtcgtatcagagcagaaagatttatacatttaatgcaaaatgaa
atcattccaaaacgtgatattattaccgaagatatgatttctgattgtattaataatgcc
ggcattgactatcaagtttttaaagaagacttgcaaaaggacaagttgactgacagcttg
aaagttgatcttcacattgcaagagaaatggaaatagaacaagctccctcacttgttttt
ttcagcgaaaatgttcatgaagaaggtttaaaagtcgaaggattatatccttatcatatt
tatacttacattattaatgagttaatgggacaacctatagagaaaaatcttcctccaaaa
ttagaatactacattcaaaagaaacaactagtaacaatggaagaacttttaacgatctat
gaatggcctgaaaaaattgctaaataaagaattaaagaaactcacacttcaacaaaaagtt
gaaaagttgcaatatccagagggagaattttggaaatctaaaatgcctcagtgttac
```

Sequence 2360
MENKSREDTNLSPVSKIEIYSFFDPFSKDCFKLSAILSKLRIEYNKYIKVRHILNPSLKV
LTKCQAQSTSDFDNIALAYKAAELQGRIRAERFIHLMQNEIIPKRDIITEDMISDCINNA
GIDYQVFKEDLQKDKLTDSLKVDLHIAREMEIEQAPSLVFFSENVHEEGLKVEGLYPYHI
YTYIINELMGQPIEKNLPPKLEYYIQKKQLVTMEELLTIYEWPEKLLNKELKKLTLQQKV
EKLQYPEGEFWKSKMPQC*

Sequence 2361
Contig_0745_pos_5858_4050,

is similar to (with p-value 0.0e+00)
>gp:gp!D88209|D88209_1 Bacillus licheniformis DNA for Pi-pep
tidase, complete cds. NID: g1651215.
atgagtcaacaattaacaagagaagaacaggaacgtaaatatcctgaatatacatggat
5 ttaacaactattttttaaaagtgatgaagcatttgaagaagcttttaaaagtattgaagct
aaaataggtgaagaagaaaaatttaaaggtcatcttggtgaaagtgctgaaacattatat
gaagcgctaagtcttgaagacgagttaggtacaaaattagaaaaggtatatgtatacgca
catttaaaacaagatcaagatactgcaaatgataaatataccggtttagaagcgcgtgca
catcaacttgttattaaatatagctctgcatggagttttttagtacctgaaattttacaa
10 ctagatgaagctactattcaatcttttatcgattctaatgatgatttaaaacgatatgaa
ttcgatttgaaattgattaatgagaaacgtccacatatattagatgcgaatacagaaaag
ttattaacagaagcacaagacgcactttcaacgccttctaatgtatatggaatgttcagc
aatgcagatttagaatttgaagatgctatagataaagatggtcaagcttatcctttaaca
caaggtacatttatcaagtatttagagtctgatgatcgtgagttaagagcttctgctttt
15 agaaatgtttataaagcatacggtgcgcataataacacgctaggtgctactttagctggt
gaggttaagaaaaatgtatttaatgctagaactcatcattatcgttcagcacgtgaaaga
gctttaagtaataatcatattccagaagctgtttacgataacttaatcaaaacggtccat
aaatacttacctttattacacagatacacgaagcttagacaagagttactaggtttagac
gatttaaaaatgtatgatctttatacacctcttgttaaagatgtcaaatttgaaatgcca
20 tatgaagaggcaaaatcctggatgttaaaagcacttgagccaatgggagaagaatactta
aacgtggttaaggaaggtctagataaccgttgggtcgatgtatatgaaaataaaggtaaa
cgttcaggcggatattcatccggtggacatttaactaatcctttcattttacttaactgg
tcagacactgtttctgatttatatactttagtacatgaatttggtcactctgcacatagt
tactttagtagacagaatcaaccatcaaatttaagcgattatacaatctttgtcgctgag
25 gtagcatcaacttgtaatgaggctttacttagtgactacatggacaaacatttagatgat
gaacgacgtctattgttacttaaccaagaattagaacgatttagagcaacactattccgt
caaacaatgtttgctgaatttgaacataaaatacatcaaatagaagaagctggggagccg
ttaacgccaaatcgtatgaatgaagaatatgctaaactgaacaaactatattttggtgaa
gcagtagaaactgacgatgatattagtaaagaatggtcacgtattcctcatttctatatg
30 aattattatgtatatcaatacgcaactggttatagtgcagctcaaagtttaagtcatcaa
attttaactgagggtcaacctgctgttgaacgatatatcaatgaattcttaaaaaagggt
agctcaaactatccgattgaaattttaaaaaatgcaggtgttgacatgacaacacctcaa
ccaatagaggaagcttgtgaagtattcgaacaaaaattagatgctttttgaaaagttaatg
aaagcttag
35
Sequence 2362
MSQQLTREEQERKYPEYTWDLTTIFKSDEAFEEAFKSIEAKIGEEEKFKGHLGESAETLY
EALSLEDELGTKLEKVYVYAHLKQDQDTANDKYTGLEARAHQLVIKYSSAWSFLVPEILQ
LDEATIQSFIDSNDDLKRYEFDLKLINEKRPHILDANTEKLLTEAQDALSTPSNVYGMFS
40 NADLEFEDAIDKDGQAYPLTQGTFIKYLESDDRELRASAFRNVYKAYGAHNNTLGATLAG
EVIKRVFMARTHHYRSARERALSNNHIPEAVYDNLIKTVHKYLPLLHRYTKLRQEILGLD
DLKMIDLYTPLVKDVKFEMPYEEAKSWMLKALEPMGEEYLNVVKEGLDNRWVDVYINKGK
RSGGYSSGGHLTNPFILLNWSDTVSDLYTLVHEFGHSAHSYFSRQNQPSNLSDYTIFVAE
VASTCNEALLSDYMDKHLDDERRLLLLNQELERFRATLFRQTMFAEFEHKIHQIEEAGEP
45 LTPNRMNEEYAKLNKLYFGEAVETDDDISKEWSRIPHFYMNYYVYQYATGYSAAQSLSHQ
ILTEGQPAVERYINEFLKKGSSNYPIEILKNAGVDMTTPQPIEEACEVFEQKLDAFEKLM
KA*

Sequence 2363
50 Contig_0745_pos_1454_846,
is similar to (with p-value 2.0e-77)
>sp:sp|O06446|SECA_STAAU PREPROTEIN TRANSLOCASE SECA SUBUNIT
. >gp:gp|U97062|SAU97062_1 Staphylococcus aureus NCTC 8325 S
ecA (secA) gene, complete cds. NID: g2078389.
55 atgggtggtattgctatacataaaggtgatattgcagaaatgagaacaggtgaagggaaa
acattgactgcaaccatgccgacgtatttgaatgctttagctggtagaggtgtacatgtt
attacagtcaatgaatatctatcaagttcacaaagtgaagaaatggctgaactatataac
tatcttggcttaactgtaggtttgaacttaaatagtaagtcaactgaagaaaaacgtgag
gcttacgcacaagatatcacttatagtacgaataatgaacttgggtttgattatcttaga

581

```
gataatatggtgaactatgctgaagagagagtaatgcgtcctctacattttgcaattatt
gatgaggtcgattccatattgatcgacgaagcaagaacacctttaattatttctggtgaa
gcggaaaaatctacttctttatatggaaaattgattattattagttataaattattcaaa
aataaagagcaactattttatgttgcatatgaaataaagggaaagaaagtctactataaa
cgagatataaatcctaaagattatattaaagaacataaacctgattgtaaggatatacgt
agaaaatga
```

Sequence 2364
MGGIAIHKGDIAEMRTGEGKTLTATMPTYLNALAGRGVHVITVNEYLSSSQSEEMAELYN
YLGLTVGLNLNSKSTEEKREAYAQDITYSTNNELGFDYLRDNMVNYAEERVMRPLHFAII
DEVDSILIDEARTPLIISGEAEKSTSLYGKLIIISYKLFKNKEQLFYVAYEIKGKKVYYK
RDINPKDYIKEHKPDCKDIRRK*

Sequence 2365
Contig_0745_pos_0_300,
putative peptide of unknown function
```
atggattattattatcaatttcataagcatcaacattacttcctatgtcatgatattta
gaagatgcttggaaatcacaaaatacttttcaaaacatgatgcagtagtaagtctcatt
ttgtttgcaacagctaattatcactatcgtcgtgggaatttaaaaggcgcattaaaatct
tttgaaaaagctttattaactattaaccacgctaaagatgacactcaattacaccttaat
atcaatgaatttaaacaacttattgtgaaaatgattgaagcagtaaaattacaaaaaACT
```

Sequence 2366
MDYYYQFHKHQHYFLCHDILEDAWKSQNTFSKHDAVVSLILFATANYHYRRGNLKGALKS
FEKALLTINHAKDDTQLHLNINEFKQLIVKMIEAVKLQKT

Sequence 2367
Contig_0746_pos_541_1500,
is similar to (with p-value 0.0e+00)
>sp:sp|P05425|ATKB_ENTFA POTASSIUM/COPPER-TRANSPORTING ATPAS
E B (EC 3.6.1.36). >pir:pir|B45995|B45995 Cu2+-transporting
ATPase (EC 3.6.1.-) - Enterococcus hirae >gp:gp|L13292|ENECO
PPUMP_2 Enterococcus hirae ATPase (copA) gene, complete cds;
ATPase (copB) gene, complete cds. NID: g290641.
```
atgcatcatgataaccatgcctcacatcatcatagtggccatgcacatcatcatggaaat
tttaaagttaagttttttgtttcattaatttttgcaatacctatcattctcttatcgcca
atgatgggtgttaacttacctttttcaattcacatttccaggttctgaatgggtagtgtta
atattaagtacaattttattctttttatggtggtaaaccgttcttgtctggtggtaaagat
gaaattgctacaaaaaaaccaggcatgatgaccttagttgctctaggtatttcagtagct
tatatttatagcttgtatgcttttttatatgaataactttagtagtgcaactggtcataca
atggactttttttgggaattagcaaccttaatttttaattatgctattaggacattggata
gaaatgaatgctgtcggaaatgctggagatgctttaaagaaaatggcagaactattacct
aatagtgctattaaagttatggataatggccaacgcgaagaagttaaaatatcagacatc
atgactgatgatatcgtcgaagtaaaagccggagaaagcattccaacagatggtattatc
gttcaaggacaaacatcatatagatgaatccctagtcactggagaatctaaaaaagtacaa
aaaaatcaaaatgacaacgtcatcggggggttctattaatgggtctggaacaatacaagtc
aaggttacagctgttggagaagatggatatctttctcaagttatgggacttgttaatcaa
gcacaaaatgataaatctagtgctgaattgttatctgataaagtagcgggttatttattc
tactttgctgtaagtgttggcgtgatttcttttattgtctggatgctcattcaaaatgat
gttgattttgcattagaacgtcttgtaactgtgttagtcattgcttgtcccatgctttag
```

Sequence 2368
MHHDNHASHHHSGHAHHHGNFKVKFFVSLIFAIPIILLSPMMGVNLPFQFTFPGSEWVVL
ILSTILFFYGGKPFLSGGKDEIATKKPGMMTLVALGISVAYIYSLYAFYMNNFSSATGHT
MDFFWELATLILIMLLGHWIEMNAVGNAGDALKKMAELLPNSAIKVMDNGQREEVKISDI
MTDDIVEVKAGESIPTDGIIVQGQTSIDESLVTGESKKVQKNQNDNVIGGSINGSGTIQV
KVTAVGEDGYLSQVMGLVNQAQNDKSSAELLSDKVAGYLFYFAVSVGVISFIVWMLIQND

582

VDFALERLVTVLVIACPML*

Sequence 2369
Contig_0746_pos_1614_2543,
is similar to (with p-value 2.0e-78)
>sp:sp|P05425|ATKB_ENTFA POTASSIUM/COPPER-TRANSPORTING ATPAS
E B (EC 3.6.1.36). >pir:pir|B45995|B45995 Cu2+-transporting
ATPase (EC 3.6.1.-) - Enterococcus hirae >gp:gp|L13292|ENECO
PPUMP_2 Enterococcus hirae ATPase (copA) gene, complete cds;
ATPase (copB) gene, complete cds. NID: g290641.
atggataaaactggtactttaactgagggtaacttttctgtgaatcattatgagagcttt
aaaaatgatttgagtaatgatacaatattaagccttttcgcctcattagaaagtcaatct
aatcacccattagctataagtattgttgattttgcgaaaagtaaaaatgtttcatttact
aatccacaagacgttaataatattccaggtgtcggattagaaggtctaattgataataaa
acatataaaataacaaatgtctcttatcttgataaacataaacttaattatgacgatgac
ttgtttactaaattagctcaacaaggtaattcaatcagttatttaattgaggatcaacaa
gtcattggcatgattgctcaaggagatcaaattaaagaaagctcaaaacaaatgatagct
gatttactatcaagaaatattacaccagtcatgcttacaggtgacaataatgaagtggca
cacgctgtcgcaaaagaattaggtattagtgatgttcacgcacaactcatgccagaagat
aaggaaagcattataaaagattatcaaagtgacggtaataaagtcatgatggtcggagac
ggtatcaacgatgcgccgagtcttataagagccgatattggtatagcaattggtgcaggc
acagatgttgcagtggattcaggtgatatcatacttgttaaaagtaatccatcagatatc
attcatttcttgactctttcaaataatactatgagaaaaatggtgcaaaacttatggtgg
ggtgcaggttataatattgttgctgtacctttagcagctggcgcattagctttat.cggg
ttaatattatcaccagctgtaggagcaatattaatgtctttaagtacagttatagtagcg
attaatgcttttacattaaaattaaaataa

Sequence 2370
MDKTGTLTEGNFSVNHYESFKNDLSNDTILSLFASLESQSNHPLAISIVDFAKSKNVSFT
NPQDVNNIPGVGLEGLIDNKTYKITNVSYLDKHKLNYDDDLFTKLAQQGNSISYLIEDQQ
VIGMIAQGDQIKESSKQMIADLLSRNITPVMLTGDNNEVAHAVAKELGISDVHAQLMPED
KESIIKDYQSDGNKVMMVGDGINDAPSLIRADIGIAIGAGTDVAVDSGDIILVKSNPSDI
IHFLTLSNNTMRKMVQNLWWGAGYNIVAVPLAAGALAFIGLILSPAVGAILMSLSTVIVA
INAFTLKLK*

Sequence 2371
Contig_0746_pos_4011_4493,
putative peptide of unknown function
atgatgaaaaaagataaagacactaatgaccaaaaaagtgagagccatatgaagcataat
gatgaaagtaaagttcctgaagatatgacatcgactaatgagggtgaatttaaagtagga
gataaagtaacgattacagcaggacatatgccaggtatgaaaggtgcagaagctactgta
aaaggtgcgtataaaacatatgcctatgttgtaagttataaacccacaaatggaaatgaa
aaagtaaacaatcataaatggatcgtaaacgaagaaatcaaagatgcacctaaagatgga
tttagtaaaggcgatactgttaaattagaagcaagtcatatgtctggtatgaaaggtgct
acagccaatatagataacgtgaaaaagacgactgtttacgtagttgattacaaatccaaa
gataatggtaaaatcattaaaaatcataaatggatgacaggaaatgagctgaaagcacga
taa

Sequence 2372
MMKKDKDTNDQKSESHMKHNDESKVPEDMTSTNEGEFKVGDKVTITAGHMPGMKGAEATV
KGAYKTYAYVVSYKPTNGNEKVNNHKWIVNEEIKDAPKDGFSKGDTVKLEASHMSGMKGA
TANIDNVKKTTVYVVDYKSKDNGKIIKNHKWMTGNELKAR*

Sequence 2373
Contig_0746_pos_5109_4705,
is similar to (with p-value 8.0e-65)
>sp:sp|P30330|ARSC_STAAU ARSENATE REDUCTASE (ARSENICAL PUMP
MODIFIER). >pir:pir|D41903|D41903 arsenate reductase (EC 1.-
.-.-) - Staphylococcus aureus plasmid pI258 >gp:gp|M86824|PI

2ARSRBC_4 Plasmid pI258 arsenic resistance operon (arsRBC) g
enes, complete cds. NID: g150725.
gtgtcaataatgacaaagaaaataatttactttatatgtacaggtaattcatgccgtagc
caaatggctgaaggctggggcaaaaacatcttaggtgatgaatggcaagtatattctggc
ggtattgaggcacacggtgtcaatccaaaggcaattgaagcaatgaaagaagttggaatt
gatatctcaaaccatacttctaatttaattgacaacactatactaaaccaatccgattta
gtagtaactctatgtagtgatgctgacaataattgtcctatattaccgccgaatgtaaaa
aaagaacattggggatttgatgatccagcaggcaaaccttggtcagaattccaacgtgtt
agagatgaaatcaaaaccgccattgaatcatttaaaactagataa

Sequence 2374
VSIMTKKIIYFICTGNSCRSQMAEGWGKNILGDEWQVYSGGIEAHGVNPKAIEAMKEVGI
DISNHTSNLIDNTILNQSDLVVTLCSDADNNCPILPPNVKKEHWGFDDPAGKPWSEFQRV
RDEIKTAIESFKTR*

Sequence 2375
Contig_0746_pos_4270_3908,
putative peptide of unknown function
atgattgtttacttttttcatttccatttgtgggtttataacttacaacataggcatatgt
tttatacgcacctttтacagtagcttctgcacctttcatacctggcatatgtcctgctgt
aatcgttactttatctcctactttaaattcaccctcattagtcgatgtcatatcttcagg
aactttactttcatcattatgcttcatatggctctcacttttттggtcattagtgtcttt
atcttttttcatcatattaaaatctccctattgattagtttgttacttttatttgaccca
tcattccattatcttcatgctcaagtatatgacagtgaaacatgtatgttccagtatttt
taa

Sequence 2376
MIVYFFISICGFITYNIGICFIRTFYSSFCTFHTWHMSCCNRYFISYFKFTLISRCHIFR
NFTFIIMLHMALTFLVISVFIFFHHIKISLLISLLLLFDPSFHYLHAQVYDSETCMFQYF
*

Sequence 2377
Contig_0747_pos_4050_4487,
is similar to (with p-value 8.0e-27)
>sp:sp|P42435|NASD_BACSU NITRITE REDUCTASE (NAD(P)H) (EC 1.6
.6.4). >gp:gp|D30689|BACNARB_4 Bacillus subtilis DNA around
narB region (nasB operon and nasA gene). NID: g710016. >gp:g
p|Z99105|BSUB0002_159 Bacillus subtilis complete genome (sec
tion 2 of 21): from 194651 to 415810. NID: g2632457. >gp:gp|
D50453|D50453_33 Bacillus subtilis DNA for 25-36 degree regi
on containing the amyE-srfA region, complete cds. NID: g1805
369.
atgttatggaggtttgttatgggaagtttttttaatcggatgactcgaaaagagaatcct
actgtagagtctggtgttaaagattttggcgtcatatctgttgaaaatggctaccaaata
tttatcggaggtaatggtggtactgatgttactgtaggtaaattgttaacgacagttgaa
accgaagatgaagtgattcaattatgtggtgccctcatgcagtattacagagaaacaggt
gtttacgctgaaagaacagcaccatggttagaacgtatgggctttgaaaatgtcaagaat
gtcttattaaatcaagaaaagcaaaaagaactgtatttaagaattatggaagccaaaaaa
gctgttgagaatgaaccatgggaaactattgttgaaaataaagaagcacaaaaaatcttt
gaagttgagaaggtgtaa

Sequence 2378
MLWRFVMGSFFNRMTRKENPTVESGVKDFGVISVENGYQIFIGGNGGTDVTVGKLLTTVE
TEDEVIQLCGALMQYYRETGVYAERTAPWLERMGFENVKNVLLNQEKQKELYLRIMEAKK
AVENEPWETIVENKEAQKIFEVEKV*

Sequence 2379
Contig_0747_pos_4490_4804,
is similar to (with p-value 3.0e-22)

>sp:sp|P42436|NASE_BACSU ASSIMILATORY NITRITE REDUCTASE (NAD
(P)H) SMALL SUBUNIT (EC 1.6.6.4). >gp:gp|D30689|BACNARB_5 Ba
cillus subtilis DNA around narB region (nasB operon and nasA
gene). NID: g710016. >gp:gp|Z99105|BSUB0002_158 Bacillus su
btilis complete genome (section 2 of 21): from 194651 to 415
810. NID: g2632457. >gp:gp|D50453|D50453_32 Bacillus subtili
s DNA for 25-36 degree region containing the amyE-srfA regio
n, complete cds. NID: g1805369.
atgaaagctaaagaaaagattaaagttacaacaatgaatgaaatgattcctcaaataggc
aaaaaagtagttgtaaacgaaaaagaaataggtattttttctcacagataatggtgattta
tatgccattggaaatatatgtccacataaagaaggaccgttgtctgaagggactgtaagt
ggtgattatgtttactgtccgttacacgatcaaaaaatagctttaaaaactggagaagta
caacaacctgatacaggatgtgtagagacatacgaagtagaagttattgatggagatatt
tacttatgtctataa

Sequence 2380
MKAKEKIKVTTMNEMIPQIGKKVVVNEKEIGIFLTDNGDLYAIGNICPHKEGPLSEGTVS
GDYVYCPLHDQKIALKTGEVQQPDTGCVETYEVEVIDGDIYLCL*

Sequence 2381
Contig_0747_pos_3285_2647,
is similar to (with p-value 8.0e-40)
>sp:sp|P39592|YWBI_BACSU HYPOTHETICAL TRANSCRIPTIONAL REGULA
TOR IN EPR-GALK INTERGENIC REGION. >pir:pir|S39679|S39679 hy
pothetical protein - Bacillus subtilis >gp:gp|X73124|BSGENR_
25 B.subtilis genomic region (325 to 333). NID: g413923. >gp
:gp|Z99123|BSUB0020_126 Bacillus subtilis complete genome (s
ection 20 of 21): from 3798401 to 4010550. NID: g2636240.
atgaatgacattgtgaacgttcaaaaaggtcatattaaaataggcttatcaccaatgatg
aatgttcaaatgtttacaaatgcattgaatcagtttcacagactctatcctaatgtgaca
tatgaagtgattgagggtggtggtaaaattgttgagaacttaacatctaatgatgatgtg
gatattggtattactacattacctgtagatcacactgaatttcattcaacttctttatat
aatgaagaattattattagtagtaagtaatgaccatcatttagcacatttaaataaagta
gacatggcagatttgaaagatgaagagtttgtttatttcatgatgattattatttaaaa
gatcaaattatagagaactgtaaaaggctaggctattaccctaaaactgttgctaatatt
tctcaaattagtttatcgctaatatgattcaacaaggaataggaattagtatcgttcca
gaaagtttagttaatttaatggggaataacgtaacgtccattcaattagagaatgttgaa
ttatcatggcatcttggcgtgatatggagaaaagatgcttatctcaatcatgtaactcgc
aaatggattgaatttatttctgagatgaaaccaacatag

Sequence 2382
MNDIVNVQKGHIKIGLSPMMNVQMFTNALNQFHRLYPNVTYEVIEGGGKIVENLTSNDDV
DIGITTLPVDHTEFHSTSLYNEELLLVVSNDHHLAHLNKVDMADLKDEEFVLFHDDYYLK
DQIIENCKRLGYYPKTVANISQISFIANMIQQGIGISIVPESLVNLMGNNVTSIQLENVE
LSWHLGVIWRKDAYLNHVTRKWIEFISEMKPT*

Sequence 2383
Contig_0747_pos_1831_863,
is similar to (with p-value 8.0e-89)
>pir:pir|A25805|A25805 L-lactate dehydrogenase (EC 1.1.1.27)
 - Bacillus subtilis
atgaaggagttcgttaaaatgaaaaaatttgggaaaaaagttgtttttagtaggagacggt
tccgtaggttcaagttatgcatttgctatggtgactcaaggaattgcagatgaatttgta
attattgatattgcaaaagataaagtggaagcagacgttaaagatttaaaccatggtgca
ctttacagttcttcaccagtgactgtaaaagctggagaatatgaagattgtaaagatgca
gatttagttgttattacagcaggtgcacctcaaaaacgggtgaaactcgtttacaactt
gttgagaaaaatactaaaatcatgaaaagtatcgtaactagtgtcatggatagtggcttt
gatggtttcttcctaattgctgcaaacccagttgatatcttaacacgttatgttaaagaa
gttacaggtttaccagctgaacgtgttattggttctggtacagtgcttgatagtgcaaga

```
        ttcagatatttaataagtaaagaattaggtgttacatcaagtagtgttcacgctagcatt
        ataggtgaacatggtgactctgaacttgcagtttggtctcaagcaaacgttggaggtatt
        tcagtgtatgatacattgaaagaagaaactggtagcgatgctaaagcgaatgaaatttat
        attaatacaagagatgctgcttacgatatcattcaagctaaaggatctacgtattatggt
    5   atagctctagcactattacgtatttctaaagctttactaaataatgaaaatagtatttttg
        acagtttctagtcaacttaatggtcaatatggatttaacgatgtttatcttggcttacca
        acacttatcaatcaaaatggtgcagttaaaatttatgaaacaccattaaatgataacgaa
        ctacaattactagaaaaatcagtgaaaactttagaagacacttatgattctataaaacat
        ttagtttaa
   10
        Sequence 2384
        MKEFVKMKKFGKKVVLVGDGSVGSSYAFAMVTQGIADEFVIIDIAKDKVEADVKDLNHGA
        LYSSSPVTVKAGEYEDCKDADLVVITAGAPQKPGETRLQLVEKNTKIMKSIVTSVMDSGF
        DGFFLIAANPVDILTRYVKEVTGLPAERVIGSGTVLDSARFRYLISKELGVTSSSVHASI
   15   IGEHGDSELAVWSQANVGGISVYDTLKEETGSDAKANEIYINTRDAAYDIIQAKGSTYYG
        IALALLRISKALLNNENSILTVSSQLNGQYGFNDVYLGLPTLINQNGAVKIYETPLNDNE
        LQLLEKSVKTLEDTYDSIKHLV*
```

```
        Sequence 2385
   20   Contig_0747_pos_800_450,
        is similar to (with p-value 1.0e-34)
        >gp:gp|L16975|LACALS_2 Lactococcus lactis alpha-acetolactate
         synthase (als) gene, complete cds. NID: g473900. >gp:gp|A23
         961|A23961_1 L. lactis alpha-acetolactate synthase gene. NID
   25    : g809617.
        atggcggaaaaacaatattctgcagcacaaatggtaattgatactttaaaaaataatgga
        gttgagtatgtatttggtattccaggtgcgaaaatcgactacttatttaatgcactagag
        gatgacgatattgaattagtcgttacgcgtcatgaacaaacgcagcgatgattgcacaa
        ggtattggtcgtttaacaggaaaaccaggtgtggctattactacaagtggcccagggta
   30   agtaacttaactactggtttattaactgcaacttctgaaggtgaccctgtattagctatc
        ggtggtcaagttaaaagaaatgacttattacgtttaacatcaatacgctaa
```

```
        Sequence 2386
        MAEKQYSAAQMVIDTLKNNGVEYVFGIPGAKIDYLFNALEDDDIELVVTRHEQNAAMIAQ
   35   GIGRLTGKPGVAITTSGPGVSNLTTGLLTATSEGDPVLAIGGQVKRNDLLRLTSIR*
```

```
        Sequence 2387
        Contig_0748_pos_1246_3342,
        is similar to (with p-value 0.0e+00)
   40   >sp:sp|P52026|DPO1_BACST DNA POLYMERASE I (EC 2.7.7.7) (POL
        I). >gp:gp|L42111|BACPOL_1 Bacillus stearothermophilus DNA p
        olymerase I (pol) gene, complete cds. NID: g806280.
        atgaaaggtctaatggggggatacctctgacaatattcctggcgttgctggtgtcggcgaa
        aagacggctattaaattacttaatcaatttgagtcagtagaaggggtctatgaacatatt
   45   gaggaggtcactgcaaaaaaattaaaagaaaaactcatcaatagtaaagatgatgcctta
        atgagtaaagatttagcaacaatcaatgttcacagtccgattgaagtatcattagaagat
        acaaaattaactctacaagacgacactacagaaaaaattgaactatttaaaaagctagaa
        tttaaacaactattagcagatatagacacatcctctacgaatgaagaagtcatagataaa
        actttgaaattgagcaagactttcaaaatgtagatttgaatgatttaaacgaagcggta
   50   atacattttgaactcgaaggcactaattatcttaaagcacactattctcaagtttggtttt
        tatacaaatcatcaacatgtagtgataaatgctgaggatgtaaaggattataaacattta
        gttcaatggcttgaagataaaaatacaactaaaattgtctatgatgcaaaaaaaacttat
        gtatctgctcatcgattagggattaatatagaaaatattgaatttgatgttatgttagca
        agctatattattgacccatcacgttctattgatgacgttaaatctgtggtaagtttatat
   55   ggacaaaattatgtaaaagataatattacaatatttgggaaaggtaagaaaacatcatata
        cctgaagaaccaattctaaacgaacacattgcctctgtgacagaagctatagcagctgta
        actccaaccatgaaatcacagttagaagattataatcaaattgaactgttgaaagattta
        gaattaccattagcaagaattttaagtgaaatggaagaaattggtatatataccgtatc
        aatgatttgaaagaaatggaattcgaaattcaaaaaaaattggatgtattaatatccaat
```

```
      attcatgagtcggctggtgaagcgtttaatatcaattctcctaagcaattaggtgttgtt
      ttatttgaaacattacaattgcctgtcattaagaagaccaaaacgggctattcaacagct
      gtagacgtattagaaaaactacaaggtgagcatcctattatagatgatattttagaatat
      agacaacttgctaagttgcaatctacgtatgtagagggattacaaaaagtaataagcaaa
  5   gatcatagaattcacacacgttttaatcaaacgcttgctcaaactggtagattatcaagt
      atagatcctaatttacaaaatatacctatacgattagaagaaggaagaaagattagaaag
      gcctttaaaccaacttctaaagatagtgtgattttatctgctgattattcacaaattgag
      ttacgtgtacttgctcatattacgcaagatgaaagtttaaaacatgcatttataaacgga
      catgatattcacactgcaacagcaatgaaagtatttaatgttgaatctgaccaggttgat
 10   agtttaatgagacgtcaagcaaaagctgttaactttggtattgtatatggtatcagcgat
      tatggattgagtcagagcttgggtattactagaaaacaagcaaaagcatttattgatgat
      tatttagctagttttccaggtgtaaaacaatatatgtcagacattgttaaagatgcaaaa
      gcacaaggttatgtggaaacactacttcatcgtcgtcgatacattcctgatataacaagt
      agaaacgttaatttaagaagtttttgcagaaagaacagcaatgaatacacccatacaaggt
 15   agtgcagctgacataataaaaattagcaatggttaaattcagtgaaaagattaaagaaact
      aaatatcatgctaagttattattacaagttcatgatgaactcatatttgaaataccaaaa
      tcagaagtagaagatttagtaaatttgtagaagaaattatggaacaagcattagtgctc
      gatgtacctttaaaagtagattcgaattatggtgcaacatggtacgatgctaaataa
```

Sequence 2388
MKGLMGDTSDNIPGVAGVGEKTAIKLLNQFESVEGVYEHIEEVTAKKLKEKLINSKDDAL
MSKDLATINVHSPIEVSLEDTKLTLQDDTTEKIELFKKLEFKQLLADIDTSSTNEEVIDK
TFEIEQDFQNVDLNDLNEAVIHFELEGTNYLKDTILKFGFYTNHQHVVINAEDVKDYKHL
VQWLEDKNTTKIVYDAKKTYVSAHRLGINIENIEFDVMLASYIIDPSRSIDDVKSVVSLY
GQNYVKDNITIFGKGKKHHIPEEPILNEHIASVTEAIAAVTPTMKSQLEDYNQIELLKDL
ELPLARILSEMEEIGIYTDINDLKEMEFEIQKKLDVLISNIHESAGEAFNINSPKQLGVV
LFETLQLPVIKKTKTGYSTAVDVLEKLQGEHPIIDDILEYRQLAKLQSTYVEGLQKVISK
DHRIHTRFNQTLAQTGRLSSIDPNLQNIPIRLEEGRKIRKAFKPTSKDSVILSADYSQIE
LRVLAHITQDESLKHAFINGHDIHTATAMKVFNVESDQVDSLMRRQAKAVNFGIVYGISD
YGLSQSLGITRKQAKAFIDDYLASFPGVKQYMSDIVKDAKAQGYVETLLHRRRYIFDITS
RNVNLRSFAERTAMNTPIQGSAADIIKLAMVKFSEKIKETKYHAKLLLQVHDELIFEIPK
SEVEDFSKFVEEIMEQALVLDVPLKVDSNYGATWYDAK*

Sequence 2389
Contig_0748_pos_4255_4857,
is similar to (with p-value 2.0e-31)
>sp:sp|Q55515|Y553_SYNY3 HYPOTHETICAL 22.5 KD PROTEIN SLR055
3. >gp:gp|D64006|SYCSLLLH_95 Synechocystis sp. PCC6803 compl
ete genome, 25/27, 3138604-3270709. NID: g1001291.
```
      gtgattgggataactggtggtattgccactggaaaatcaacagtttcagaattattaaca
      gcatatgggtttaaaatcgtagatgctgatattgcttcacgcgaagcagttaaaaaaggc
      tctaagggtcttgaacaagttaaagagattttgggggaagaagcaattgacgaaaatggt
      gagatgaatcgtcaatatgtaggagagatagtttttaatcatcctgacttacgcgaggct
      cttaatgaaatagttcatcctattgtaagagagataatggaacaagagaaaaacaattat
      ctagaacatggatatcatgtaattatggatatcccattgttgtacgaaaatgaactacaa
      gatactgtagatgaagtttgggtggtttatacatctgaaagtattcaaatcgatcgttta
      atggagaggaataatttatcattagaagatgctaaagcacgtgtttatagtcaaatatct
      atagataaaaaaagtaggatggcagatcatgtgatagataatctaggtgataaattagaa
      cttaaacagaatttacaaaaattacttgaagaagaagggtatattcaatcggagagtgaa
      tag
```

Sequence 2390
VIGITGGIATGKSTVSELLTAYGFKIVDADIASREAVKKGSKGLEQVKEIFGEEAIDENG
EMNRQYVGEIVFNHPDLREALNEIVHPIVREIMEQEKNNYLEHGYHVIMDIPLLYENELQ
DTVDEVWVVYTSESIQIDRLMERNNLSLEDAKARVYSQISIDKKSRMADHVIDNLGDKLE
LKQNLQKLLEEEGYIQSESE*

Sequence 2391
Contig_0751_pos_2197_2541,

putative peptide of unknown function
atggagatgatagaagaacgtaatttatcagggcttattcaaacactaactttcaatcat
cccatcattcaaattcttaaagagaacacattaaatcaacttaaaatactctctcattat
ttaccagagcgacaccctgcaatggtggcaattcaatcttggtcacaatggtttactgat

5   catgggattactgaaatccaccttgatgtaactgcacaagcgcctagatcttattacaaa
ggtattttTataaaatgtcatcttaaaaatactgctcatagcgttttgacaggtggatat
tatcacggttcactagaaggttttggtttaggattaacactttaa

Sequence 2392
10   MEMIEERNLSGLIQTLTFNHPIIQILKENTLNQLKILSHYLPERHPAMVAIQSWSQWFTD
HGITEIHLDVTAQAPRSYYKGIFIKCHLKNTAHSVLTGGYYHGSLEGFGLGLTL*

Sequence 2393
Contig_0751_pos_2566_3180,
15   is similar to (with p-value 5.0e-27)
>sp:sp|Q02129|HIS1_LACLA ATP PHOSPHORIBOSYLTRANSFERASE (EC 2
.4.2.17). >pir:pir|D45734|D45734 HisG - Lactococcus lactis s
ubsp. lactis >gp:gp|U92974|LLU92974_4 Lactococcus lactis unk
nown gene, partial cds, and HisC (hisC), unknown, HisG (hisG
20   ), unknown, HisB (hisB), unknown, HisH (hish), HisA (hisA),
HisF (hisF), HisIE (hisIE), unknown, unknown, LeuA (leuA), L
euB (leuB), LeuC (leuC), LeuD (leuD), unknown, IlvD (ilvD),
IlvB (ilvB), IlvN, IlvC (ilvC), IlvA (ilvA), AldB (aldB) and
aldR (aldR) genes, complete cds. NID: g2565137.
25   atgttacgagttgcattagcaaagggtcgtttattaaagagtttTatcgaatatttacaa
caagttaatcagatagatattgcaactgtactTttaaatagacagcgacagttattgctt
acagtcgacaacattgaaatgatTttagttaaaggaagcgatgtgcctacttatgtagaa
caaggtattgctgatgtaggaatagtgggaagtgatattctgaatggtcaaaaatataat
attaataaattactcgatTtgccatttggtaaatgtcattTtgcgttggcggcaaagcca
30   gaaacatctcgctataaaaaagtagcaacaagctatgtacatacagctactcaattcttt
aataaagaaggtatggatgtagaagtgattcaccttaacggttcagttgaattgtcatgt
gtagtggatatggttgatgctattgtagatattgtacaaactggttctacgcttacagct
aacgggctcgttgagaaaaagcatatcagtgaaattaacgctaagttaattacaaataaa
gaatcatatTttaagcaatcatctgaaatagagagactaatcaagcagttaggagtgtct
35   attaactatgcttag

Sequence 2394
MLRVALAKGRLLKSFIEYLQQVNQIDIATVLLNRQRQLLLTVDNIEMILVKGSDVPTYVE
QGIADVGIVGSDILNGQKYNINKLLDLPFGKCHFALAAKPETSRYKKVATSYVHTATQFF
40   NKEGMDVEVIHLNGSVELSCVVDMVDAIVDIVQTGSTLTANGLVEKKHISEINAKLITNK
ESYFKQSSEIERLIKQLGVSINYA*

Sequence 2395
Contig_0751_pos_3320_4003,
45   is similar to (with p-value 2.0e-41)
>sp:sp|Q02136|HISX_LACLA HISTIDINOL DEHYDROGENASE (EC 1.1.1.
23) (HDH). >pir:pir|E45734|E45734 HisD - Lactococcus lactis
subsp. lactis
gtggaaactgagaagcttgaattagagcaaagccaactaaaaaatgcatacgacatgcta
50   gataatgaaacacgagatgcattagagcaaagctatcagagaattaaagtgtaccaagaa
aatattaaggtaaaacaggaatcatctcaacaaactgaatgttatgaacgataccatcct
atcgaacgtgtaggtatttatgtgccgggaggtaaggctagctatccgtctacagtatta
atgactgcaacacttgctcaagtagcaggtgttaatgagattactgttgttaccccacct
caaaatagcggtatatgtcaagaggtgttagccgcttgttacattacaggcgttcatcat
55   gtttatcaagtcggtggagcacaaagtattgcggcgctaacttatggcacggaaactata
aaaaaagtcgacaaaatcgtaggtccagggaatcaatatgttgcttatgccaaaaagttt
gtattcggtcaagtaggcatagaccaaatcgcagaaccgacagaaatagccttgattata
gacgaaagtgctgacttagacgcaatcgcttatgacgtatttgcacaagcagaacatgat
gaaatggcttgtacttatgtgattagtgaaaatgaaaaggtacttaatcaattgaacact

ataatacaagagaaacttcagtag

Sequence 2396
VETEKLELEQSQLKNAYDMLDNETRDALEQSYQRIKVYQENIKVKQESSQQTECYERYHP
IERVGIYVPGGKASYPSTVLMTATLAQVAGVNEITVVTPPQNSGICQEVLAACYITGVHH
VYQVGGAQSIAALTYGTETIKKVDKIVGPGNQYVAYAKKFVFGQVGIDQIAEPTEIALII
DESADLDAIAYDVFAQAEHDEMACTYVISENEKVLNQLNTIIQEKLQ*

Sequence 2397
Contig_0751_pos_832_338,
is similar to (with p-value 3.0e-78)
>sp:sp|P51065|PPCK_STAAU PHOSPHOENOLPYRUVATE CARBOXYKINASE (
ATP) (EC 4.1.1.49). >gp:gp|U51133|SAU51133_1 Staphylococcus
aureus phosphoenolpyruvate carboxykinase (pcka) gene, comple
te cds. NID: g1255261. >gp:gp|L42943|STAPEPCK_1 Staphylococc
us aureus (clone KIN50) phosphoenolpyruvate carboxykinase (p
ckA) gene, complete cds. NID: g860731.
atgtatcatttcttaagtggattcacgtctaaactagctggaacagaacgtggtgttact
gaacctcaaccttcgttttcaacttgctttggtgcaccattcttacctttgagtccaaca
aagtacgctgatctacttggaaatttaatcgatattcatgatgtagatgtatatctagta
aatactggatggacaggtggtaaatatggtgtagggcgaagaattagtctacactatact
cgtgaaatggtagatcaagcaatatcaggtaaattaaaaaaatactaaatatattaaagat
gatacatttggcttaaatattccagttcaaattgacagtgtacctacaactattctgaat
cctatcaatgcttggaacaataaagataactacaaagcacaagcttacgatttgattcaa
cgctttaataataattttaaaaaattcggcaaggaagtcgaacatattgccaacaaggt
gcatttaatcaataa

Sequence 2398
MYHFLSGFTSKLAGTERGVTEPQPSFSTCFGAPFLPLSPTKYADLLGNLIDIHDVDVYLV
NTGWTGGKYGVGRRISLHYTREMVDQAISGKLKNTKYIKDDTFGLNIPVQIDSVPTTILN
PINAWNNKDNYKAQAYDLIQRFNNNFKKFGKEVEHIANKGAFNQ*

Sequence 2399
Contig_0752_pos_2931_2140,
putative peptide of unknown function
atgagcaataacatgtggaggggagaaagttgaacttaaaattagaccatatcattcac
tatttacatcaattagagtcatttaagtttcccggagaaatattagaattgcaaaatggt
ggaagacatcatcatttgggcacctttaatcaaatagcaccgattaaaaatagttatatc
gaattgctagatgttgaaaatgagtcaaaacttagcaatatagctaaaactgaagaaggt
cgtgtatcatttgctacaaaaatagtgcaggatcattttaaacaaggggttaaaggtatt
tgttttagaacaaaggatatataatcaggttaaaagtactttagaaaatagaggcgttgat
gtgataggtcctattgatatggaaagagaaaacaaaaaaggtcatcaaattcgttggaga
ttgctatatattgctaaccctgactatacagtcaaaccacctttctttatagaatgggat
aacaacaaaaagcaaaacctatcacaaatacataatttcaacttgtcatcgtttaaaatt
aaagaggtgattattactagcactcaacgtgaaacaacagtaagtctttggaaagaatgg
tataacctgaaaatagtaaatgaaacggctacatctactgatctcaaattagaaactgat
gaagttatctataaaatagaagacggcaaagattcaggttttcatacattaataatgacc
gatatcaatgccacagcaccatattcaatatttatacgtggtgctaaatatcgttttgag
ccacccaactag

Sequence 2400
MSHNMGVRCRKLNLKLDHIIHYLHQLESFKFPGEILELQNGGRHHHLGTFNQIAPIKNSYI
ELLDVENESKLSNIAKTEEGRVSFATKIVQDHFKQGVKGICFRTKDINQVKSTLENRGVD
VIGPIDMERENKKGHQIRWRLLYIANPDYTVKPPFFIEWDNNKKQNLSQIHNFNLSSFKI
KEVIITSTQRETTVSLWKEWYNLKIVNETATSTDLKLETDEVIYKIEDGKDSGFHTLIMT
DINATAPYSIFIRGAKYRFEPPN*

Sequence 2401
Contig_0752_pos_1740_490,

589

is similar to (with p-value 3.0e-24)
>gp:gp|AF099966|AF099966_1 Staphylococcus sciuri factor esse
ntial for methicillin resistance FEMA (femA) gene, complete
cds. NID: g3820631.
5   atggaaaagatgaacatcactaatcaacaacatgacgcatttgtgaaatctcatcccaat
    ggtgatttattacaattatctaagtgggcagatacgaaaaaattaacaggatggtattca
    agaagaattgctgtcggtgaaaatggtcaaattaaaggtgttggccagctactattcaaa
    aaaatacctaaacttccatacactttatgctatgtatctaggggatttgtagctgattat
    aataataaagaagtgttagaagctctacttagctatgctaaagaagtagcaaaagatgaa
10 aagtcgtatgctatcaaaatagatcccgatgtcgaagtagataaaggtgcagaagcactt
    aaaaatctacgtgagcttgggtttaaacataaaggttttaaagaaggactgtctaaagac
    tatattcaaccaagaatgactatgattacgcctattgacaaaacagatgatgaattagtt
    caaagtttcgaacgtcgaaatcgttcaaaagtaagacttgcactgaagcgtggaactaaa
    gtagaacgatcaaatcgcgagggcttaaaatctttgctaatttaatgaagataactggg
15 gagagagatggtttttttaactcgagatattagttattttgaaaatatatatgatgcactt
    catgaagacggtgatgcagaactcttccttgttaaattagagcctaagccagtattagat
    acggttaatcaagatcttgaagcacaattagctgagaaagagaaattacaatcaaaaaag
    caagataaaaagacacttaataaacttaatgatattgataataaaattaagaaaacaaat
    gaattaaaatcggatttaacagaacttgaaaaaagcgagccagaaggtatttacttgtca
20 ggagcgctcttaatgtttgcaggaaacaaatcttactatctctatggcgcttcctcgaat
    gactatcgtgatttcttaccaaaccatcacatgcaatttgaaatgatgaaatatgcacgt
    gagcatggtgcaacaacctatgactttggtggtacagataatgatcctgataaagattca
    gaacattatgggttgtgggcttttaaacgagtttggggtacatatttaagtgaaaaaatt
    ggagaatttgattatgtattaaatcaaccgctatatcatttagttgagaaagtgaaacct
25 cgtttaacgaaagctaaaattaaaatatcacgtaaacttaaaggtaaataa

Sequence 2402
MEKMNITNQQHDAFVKSHPNGDLLQLSKWADTKKLTGWYSRRIAVGENGQIKGVGQLLFK
KIPKLPYTLCYVSRGFVADYNNKEVLEALLSYAKEVAKDEKSYAIKIDPDVEVDKGAEAL
30 KNLRELGFKHKGFKEGLSKDYIQPRMTMITPIDKTDDELVQSFERRNRSKVRLALKRGTK
VERSNREGLKIFANLMKITGERDGFLTRDISYFENIYDALHEDGDAELFLVKLEPKPVLD
TVNQDLEAQLAEKEKLQSKKQDKKTLNKLNDIDNKIKKTNELKSDLTELEKSEPEGIYLS
GALLMFAGNKSYYLYGASSNDYRDFLPNHHMQFEMMKYAREHGATTYDFGGTDNDPDKDS
EHYGLWAFKRVWGTYLSEKIGEFDYVLNQPLYHLVEKVKPRLTKAKIKISRKLKGK*

35

Sequence 2403
Contig_0752_pos_0_359,
putative peptide of unknown function
atggtagttttaattatattaggtggcgtttattcaagcgccaaattaaaacttgaatta
40 ttaccagatgttgaaaatccagttatttcagttcaaactacaatgtctggagcaacaccc
cagtcaacacaagatgaaataagtagcaagattgataatcaagtacgctcgttggcctac
gtaaatagtgtacagactgaatctatacctaatgcttctatagtaactgtagaatacgat
aatggtacagatatggataaagctgaagaacaattaaaaaaagaaatcgacaaaattaag
tttaaagatggcgttggtgaacccgaattaacaaggaactctatggatgcttttTGTTTG

45

Sequence 2404
MVVLIILGGVYSSAKLKLELLPDVENPVISVQTTMSGATPQSTQDEISSKIDNQVRSLAY
VNSVQTESIPNASIVTVEYDNGTDMDKAEEQLKKEIDKIKFKDGVGEPELTRNSMDAFVX

50

Sequence 2405
Contig_0753_pos_2155_2514,
putative peptide of unknown function
atgataaatattatattgaagaaaatagacttggaggtaattagaatgttcgttgttaca
55 aatagaatcactgtaaaaaaaggatatgcaaaacaaatggcgcctaattttactaaagga
ggacctattgaatctttaaagggctttgaaggtattgaagtttggcaaattgataaagat
gattatagcgaagatatgtatgtaaatagttggtgggaaactgaagaagattttaaaaat
tgggtgaatagtgatgtatttaaacaagcacataaaaatactggaaaatccgaagattca
ccagtcattaaaagcgaaattgttaaatcaaatgtttttatcttctttgaacagaagataa

Sequence 2406
MINIILKKIDLEVIRMFVVTNRITVKKGYAKQMAPNFTKGGPIESLKGFEGIEVWQIDKD
DYSEDMYVNSWWETEEDFKNWVNSDVFKQAHKNTGKSEDSPVIKSEIVKSNVLSSLNRR*

Sequence 2407
Contig_0753_pos_4019_2700,
is similar to (with p-value 3.0e-88)
>sp:sp|P30267|YKAA_BACFI HYPOTHETICAL 50.9 KD PROTEIN IN KAT
A 3'REGION (ORF A). >pir:pir|S27491|S27491 hypothetical prot
ein A - Bacillus firmus >gp:gp|L02548|BACKATA2_1 B.firmus OR
F A and ORF B, complete cds. NID: g143118.
atgtcattgttagggattatattattttttagttcctataccagtcgttcaagatgcgaaag
cagcaaacgacacttcctatagctttttttagctggtttattaaaagattggcttgctggt
atcatgccaattttaattgtaaccatcataactgtatcaggtattttaacaatattatgc
tctacaatttataaaaataaattaaatcctcaaggtttaatgagcagtgctttcaacgtt
aaaataggatggcttgttttgagagtattagctgtcttcttttcttggttaacattttta
aatattggacctgaaatgattaaatctgaagatacaggtggattagtattttcaagttta
ttacctactcttgtagcagtattttttatttgctgcaatctttttacctttattaatggag
tatggtctattagaattacttggacccatctttagacctatcatgcgacctttgtttact
ttacctggtagatcgacagttgataatctagcttcatttataggtgatggtacagttggt
gttttaattactagtagacaatatggtgaaggatattactctagaagagaagcaacagta
atatccacaacctttagtgttgtatctattacgttcgctattgtcattgccgaaacaatt
agaatgcaagatcaatttttctattttttatttaacagttgtcatttcatgcttaattgca
gcaatgattatgccaagaatttggccacttaaaaatattcctgacgaatatgctaaagaa
gtaagtgaagaggctcgtaatgaacagctaccagaaggcaaaacagcattaaaatatggt
tttgatttagcaactgaagttggaattaaatcgccagggtttaaagaatttttaatttca
ggttttaaaacagttgtagatatgtggtttgtaattttaccagttgttatgagtatagga
acaatagctaccattattgctaactacacgcctgtttttgaaattataggaaaaccattt
gttccagtactagaattgttacaaattccagaagcacatgaagcatcacaaacaatttta
attgggtttgccgatatgttcttaccttcaattcttattgaaggggggttcaaaatgatgta
acacgttttgtaattggagcattgagtatctcacaacttgtgtatttatctgaagtaggc
ggcgtgattcttggttctaaaattccagttagtataagtaaattatttatgatttttttta
attcgtactatcattacgcttccaataattgctttattagcgcatttatttatcggataa

Sequence 2408
MSLLGIILFLVPIPVVQDGKQQTTLPIAFLAGLLKDWLGGIMPILIVTIITVSGILTILC
STIYKNKLNPQGLMSSAFNVKIGWLVLRVLAVFFSWLTFLNIGPEMIKSEDTGGLVFSSL
LPTLVAVFLFAAIFLPLLMEYGLLELLGPIFRPIMRPLFTLPGRSTVDNLASFIGDGTVG
VLITSRQYGEGYYSRREATVISTTFSVVSITFAIVIAETIRMQDQFFYFLTVVISCLIA
AMIMPRIWPLKNIPDEYAKEVSEEARNEQLPEGKTALKYGFDLATEVGIKSPGFKEFLIS
GFKTVVDMWFVILPVVMSIGTIATIIANYTPVFEIIGKPFVPVLELLQIPEAHEASQTIL
IGFADMFLPSILIEGVQNDVTRFVIGALSISQLVYLSEVGGVILGSKIPVSISKLFMIFL
IRTIITLPIIALLAHLFIG*

Sequence 2409
Contig_0753_pos_1617_1108,
putative peptide of unknown function
atgacaggtaaaacacacgcatcatgtggcttttttagtcggtgcaataaccacacaatat
tttcatacagatatatttacttctatatcagtgattgtactttcagtcatttcaagtata
ttgccagatatatgtcatacacaaagtaaaataggaagacgatttaggcttactagtttt
tttgtcagaattttatttggtcatagaacatttacgcattcacttttatttattatagga
attagttttttactgtacttcatacaaactccgatgtattatatggttgcaattgttatt
ggtatgttttcgcatgttatacttgatatattaaccccaagaggtgttaaactattatat
cctttaccatttaatatcgtatcacccattcattttaaaactgggggactagtagatgta
tctctagctactgcattaagtgttggtgcgatatatactttatttcaaccatatttaaat

actatgatgcactattggttaatcaaataa

Sequence 2410
MTGKTHASCGFLVGAITTQYFHTDIFTSISVIVLSVISSILPDICHTQSKIGRRFRLTSF
5 FVRILFGHRTFTHSLLFIIGISFLLYFIQTPMYYMVAIVIGMFSHVILDILTPRGVKLLY
PLPFNIVSPIHFKTGGLVDVSLATALSVGAIYTLFQPYLNTMMHYWLIK*

Sequence 2411
Contig_0754_pos_2035_2466,
10 putative peptide of unknown function
atgatacaaggtttaggctatttattgtccaatataacagattataaagaattaacgaat
ttagctcaaaatggagatcgtgatgccattgatttaaaagtaaaacatatttataaagat
actgaaccaccaattcctggagatttaacagcagcaaattttggaaatgtattacatcac
ttagataatcagtttacatcagctaacaaacttgcctctgcaattggcgtcgttggtgaa
15 gttataacaactatggctattacattagcacgtgaatataagactaagcacgttgtatat
atcggttcatcatttaataacaatcaattactacgtgaagttgttgaaaattacactgtt
ctaagaggatttaaaccgtactatattgagaatggtgctttttcaggcgctttaggagca
ctttacctctaa

20 Sequence 2412
MIQGLGYLLSNITDYKELTNLAQNGDRDAIDLKVKHIYKDTEPPIPGDLTAANFGNVLHH
LDNQFTSANKLASAIGVVGEVITTMAITLAREYKTKHVVYIGSSFNNNQLLREVVENYTV
LRGFKPYYIENGAFSGALGALYL*

25 Sequence 2413
Contig_0754_pos_4500_5048,
putative peptide of unknown function
atggatgattataaagaatatagaaaaagacttatcgttaaattaaaaaaaacctataggt
agagatttatataatagattatataaaaatattcaagatactttagaacctgaggtttat
30 gaaattgctcctaatactaaattaggacatttttcctggttatcagaatgtaacgttatcc
cacccacaaatgcaacaaattatatcaagaaatgaacctagttggaaacaagctttaatg
aatgtaaaaggtgtttatgttataaccgacttaagtaatggcaaattatacataggatca
gcatcaggtaatactgatggaatatggcaacgatggtcggactatgccaacatagaaaat
ctaacaggtggtaataaaattattaaatgaaattaaattagataaagggaaagattacatc
35 ataaataattttcaatattcaattttagagatttttgatacaaagactaaggtagacact
ataatcaatagagaaaattattggaagaatgtattttgcactagaaaatatggtatgaac
tttaactaa

Sequence 2414
40 MDDYKEYRKRLIVKLKKPIGRDLYNRLYKNIQDTLEPEVYEIAPNTKLGHFPGYQNVTLS
HPQMQQIISRNEPSWKQALMNVKGVYVITDLSNGKLYIGSASGNTDGIWQRWSDYANIEN
LTGGNKLLNEIKLDKGKDYIINNFQYSILEIFDTKTKVDTIINRENYWKNVFCTRKYGMN
FN*

45 Sequence 2415
Contig_0754_pos_5848_7851,
is similar to (with p-value 0.0e+00)
>sp:sp|P05425|ATKB_ENTFA POTASSIUM/COPPER-TRANSPORTING ATPAS
E B (EC 3.6.1.36). >pir:pir|B45995|B45995 Cu2+-transporting
50 ATPase (EC 3.6.1.-) - Enterococcus hirae >gp:gp|L13292|ENECO
PPUMP_2 Enterococcus hirae ATPase (copA) gene, complete cds;
 ATPase (copB) gene, complete cds. NID: g290641.
atgcatcatgataaccatgcctcacatcatcatagtggccatgcacatcatcatggaaat
tttaaagttaagttttttgtttcattaattttttgcaatacctatcattctcttatcgcca
55 atgatgggtgttaacttacctttcaattcacatttccaggttctgaatgggtagtgtta
atattaagtacaattttattctttttatggtggtaaaccgttcttgtctggtggtaaagat
gaaattgctacaaaaaaaccaggcatgatgacccttagttgctctaggtatttcagtagct
tatatttatagcttgtatgcttttttatatgaataactttagtagtgcaactggtcataca
atggactttttttgggaattagcaaccttaatttttaattatgctattaggacattggata

\

gaaatgaatgctgtcggaaatgctggagatgctttaaagaaaatggcagaactattacct
aatagtgctattaaagttatggataatggccaacgcgaagaagttaaaatatcagacatc
atgactgatgatatcgtcgaagtaaaagccggagaaagcattccaacagatggtattatc
gttcaaggacaaacatctatagatgaatccctagtcactggagaatctaaaaaagtacaa

5   aaaaatcaaaatgacaacgtcatcgggggttctattaatgggtctggaacaatacaagtc
aaggttacagctgttggagaagatggatatctttctcaagttatgggacttgttaatcaa
gcacaaaatgataaatctagtgctgaattgttatctgataaagtagcgggttatttattc
tactttgctgtaagtgttggcgtgatttcttttattgtctggatgctcattcaaaatgat
gttgattttgcattagaacgtcttgtaactgtgttagtcattgcttgtccacatgcttta

10   ggcttggcaatacctttagtcactgcacgttctacttcaattggtgcacataatggttta
attattaaaaatagagagtctgtagaaatagctcaacatatcgattatgtaatgatggat
aaaactggtacttttaactgagggtaacttttctgtgaatcattatgagagctttaaaaat
gatttgagtaatgatacaatattaagccttttcgcctcattagaaagtcaatctaatcac
ccattagctataagtattgttgattttgcgaaaagtaaaaatgtttcatttactaatcca

15   caagacgttaataatattccaggtgtcggattagaaggtctaattgataataaaacatat
aaaataacaaatgtctcttatcttgataaacataaacttaattatgacgatgacttgttt
actaaattagctcaacaaggtaattcaatcagttatttaattgaggatcaacaagtcatt
ggcatgattgctcaaggagatcaaattaaagaaagctcaaaacaaatgatagctgattta
ctatcaagaaatattacaccagtcatgcttacaggtgacaataatgaagtggcacacgct

20   gtcgcaaaagaattaggtattagtgatgttcacgcacaactcatgccagaagataaggaa
agcattataaaagattatcaaagtgacggtaataaagtcatgatggtcggagacggtatc
aacgatgcgccgagtcttataagagccgatattggtatagcaattggtgcaggcacagat
gttgcagtggattcaggtgatatcatacttgttaaaagtaatccatcagatatcattcat
ttcttgactctttcaaataatactatgagaaaaatggtgcaaacttatggtggggtgca

25   ggttataatattgttgctgtacctttagcagctggcgcattagctttatcgggttaata
ttatcaccagctgtaggagcaatattaatgtctttaagtacagttatagtagcgattaat
gctttttacattaaaattaaaataa

Sequence 2416
30   MHHDNHASHHHSGHAHHHGNFKVKFFVSLIFAIPIIILLSPMMGVNLPFQFTFPGSEWVVL
ILSTILFFYGGKPFLSGGKDEIATKKPGMMTLVALGISVAYIYSLYAFYMNNFSSATGHT
MDFFWELATLILIMLLGHWIEMNAVGNAGDALKKMAELLPNSAIKVMDNGQREEVKISDI
MTDDIVEVKAGESIPTDGIIVQGQTSIDESLVTGESKKVQKNQNDNVIGGSINGSGTIQV
KVTAVGEDGYLSQVMGLVNQAQNDKSSAELLSDKVAGYLFYFAVSVGVISFIVWMLIQND
35   VDFALERLVTVLVIACPHALGLAIPLVTARSTSIGAHNGLIIKNRESVEIAQHIDYVMMD
KTGTLTEGNFSVNHYESFKNDLSNDTILSLFASLESQSNHPLAISIVDFAKSKNVSFTNP
QDVNNIPGVGLEGLIDNKTYKITNVSYLDKHKLNYDDDLFTKLAQQGNSISYLIEDQQVI
GMIAGGDQIKESSKQMIADLLSRNITPVMLTGDNNEVAHAVAKELGISDVHAQLMPEDKE
SIIKDYQSDGNKVMMVGDGINDAPSLIRADIGIAIGAGTDVAVDSGDIILVKSNPSDIIH
40   FLTLSNNTMRKMVQNLWWGAGYNIVAVPLAAGALAFIGLILSPAVGAILMSLSTVIVAIN
AFTLKLK*

Sequence 2417
Contig_0754_pos_3900_3526,
45   putative peptide of unknown function
atggaaataggcgtaagcaaaggttttttggtgtagcaggttttgacttactagtagat
gataataatgatgtttatgcgattgatttaaactttaggcaaaacggatcaacgagtatg
ctactttagcaaaagatttaactcatggatatcataaattttacagttacttttctaat
ggagataatacaaaattctataatgctattttaaaatacgtagaattaggtgtactttat
50   ccactttcctattacgatggagattggtatggaagaatcaagttaattctagatttggc
tgcatttggcatggggaaaataaagaattaattaatcgatatgaacaacaatttatattg
gaagctggattataa

Sequence 2418
55   MEIGVSKGFFGVAGFDLLVDDNNDVYAIDLNFRQNGSTSMLLLAKDLTHGYHKFYSYFSN
GDNTKFYNAILKYVELGVLYPLSYYDGDWYGKNQVNSRFGCIWHGENKELINRYEQQFIL
EAGL*

Sequence 2419

Contig_0754_pos_3352_2687,
putative peptide of unknown function
atgtcatacaaatatgaagcattttttaaagatattttgattaatgaatatatttatttt
gcttcaaaaaataaaaaattaattagaatacaacatgagaatttgccatatattgctatg

5   tggacagacgaaaatgttgctgagtcttatttgttacatcattcaattgattacgacaaa
atcattagagcagatattgaccgttttgtaacatatgaaatggatgaaatctttgatcca
ggtgacaaagttttagttaatgtgaataatggtgaagaaggaaacattgtagatatagtt
aaaatgactgatgagttgatgtctgaattagatgatataagaatgagagagtttattaaa
ga' gtcg.aaaatatgacgaagtatacggattgacaaacaaaggtgaaaagaatt'catt

10   atgattcagatgatgaccataacaaaccacacatcatgcctgtttggagtattaagagt
agagcgcgtaaagtacgtgatcaagattttgaagaatgtgatttaatcgaaattgaaggt
gaagtctttagtgaatggttagacaagttacgcgatgataataaagcagtagcgattgat
ttgaaatcaggtgttgttggtactgttgtatcagcgcaaaaactgtcaaatgaagcaaca
ttttaa

15

Sequence 2420
MSYKYEAFFKDILINEYIYFASKNKKLIRIQHENLPYIAMWTDENVAESYLLHHSIDYDK
IIRADIDRFVTYEMDEIFDPGDKVLVNVNNGEEGNIVDIVKMTDELMSELDDIRMREFIK
DVAKYDEVYGLTNKGEKNFIMISDDDHNKPHIMPVWSIKSRARKVRDQDFEECDLIEIEG

20   EVFSEWLDKLRDDNKAVAIDLKSGVVGTVVSAQKLSNEATF*

Sequence 2421
Contig_0754_pos_1366_569,
putative peptide of unknown function

25   atgtttataaaaaaggattttgatgatattacagttcaagtatttgaagaaaaatataga
gatgcacttaaccaatttgaattaagtgaacgacaacaaatatattcttcattgcctcaa
actgttttagatgatgcattaaaagatgaaaatcgaattgctaatgtagctttaaataaa
gaaggaaaagtagtggggttcttcgttttgcatcgttattatcaacatgaaggttatgat
acac.aaa.caatgttgtttatgtacgttcattgtcagttaatgaaaagtttcaag.ccat

30   ggatatgggacaaaaatgatgatgttttaccagagtatgttcaagcattatttcctgat
tttacacatttatacttagtagtagacgctgaaaaccaaagtgcttggaacgtttatgaa
cgtgcaggttttatgcatacagctacaaaagaagaaggacctattgggaaagaaagactt
tattatttagattttagattcaaaacatgtatcttctttaaggctaaaagagggggaagtc
acatataatgatgatattcacgtgattaatttgcttaaagatgatgtaaaggtaggcttt

35   attgcactagaacaaaatgataataaaatgaatatttctgcaatcgaagttaataagaaa
aataggaatgaggggaattgcagaaagtgctttacgccaattaccaacgtatatacgtaaa
cagtttgaagacattgaagtttttatcaattactttagtgctccagcacgcactaactcat
ttaatgtatctgcaataa

40   Sequence 2422
MFIKKDFDDITVQVFEEKYRDALNQFELSERQQIYSSLPQTVLDDALKDENRIANVALNK
EGKVVGFFVLHRYYQHEGYDTPNNVVYVRSLSVNEKFQGHGYGTKMMMFLPEYVQALFPD
FTHLYLVVDAENQSAWNVYERAGFMHTATKEEGPIGKERLYYLDLDSKHVSSLRLKEGEV
TYNDDIHVINLLKDDVKVGFIALEQNDNKMNISAIEVNKKNRNEGIAESALRQLPTYIRK

45   QFEDIEVLSITLVLQHALTHLMYLQ*

Sequence 2423
Contig_0755_pos_1306_1917,
is similar to (with p-value 8.0e-42)

50   ->gp:gp|AF012285|AF012285_40 Bacillus subtilis mobA-nprE gene
       region. NID: g3282109. >gp:gp|Z99111|BSUB0008_137 Bacillus
       subtilis complete genome (section 8 of 21): from 1394791 to
       1603020. NID: g2633699.
       atgacccaatatacttttttcacctaaagattttaaagcttttgaagtcgaaggtttagac

55   caaagaatggaagcacttaatgactatgtcagacctcaacttcatcaattaggatcttat
       tttgaagaatatttcactacacaaacaggtgaaactttttatgctcacgtagctaaacac
       gcacgtagaagtgtcaatccacctatcgatacgtgggtagcttttgctcctaataaacgt
       ggttataaaatgttaccacactttcaaatcggattgtttagaaatcagcttttcattatg
       ttcggtatcatgcacgaaggtagaaataaagaagaaaaagtgaaaatatttgataaacat

tttgataaactgacatctttaccaagtgattatagtgtttctctagatcatatgaaaact
gaaaagcactatatcaaggatatgagtaatgaagagttgcatgctgctatcgatagagtt
aaaaatgttaaaaaaggtgaattttttgttgccagaacattatcaccaaccgataaaaga
ttaaaatctgataagtcttttctaaaatttgttgaggaaacttttgatgaatttttaaaa
5   ttttatcaataa

Sequence 2424
MTQYTFSPKDFKAFEVEGLDQRMEALNDYVRPQLHQLGSYFEEYFTTQTGETFYAHVAKH
ARRSVNPPIDTWVAFAPNKRGYKMLPHFQIGLFRNQLFIMFGIMHEGRNKEEKVKIFDKH
10   FDKLTSLPSDYSVSLDHMKTEKHYIKDMSNEELHAAIDRVKNVKKGEFFVARTLSPTDKR
LKSDKSFLKFVEETFDEFLKFYQ*

Sequence 2425
Contig_0755_pos_5604_4453,
15   putative peptide of unknown function
atgttaggagagcaatatacacaaattaagcgtccagcaaatcggctaactgaaaaaata
ttaggttggtttagttgggtattcttactcatattaactattgtttcaatgtttattgcg
ctcgtatcttttagtaatgatacgtcaattgccaatttagaaaacacacttaataataat
gaactcgtacaacaaattttagccaataatgatttaagtacaactcaatttgtgatttgg
20   ttacaaaatggagtttgggcaattattgtttattttattgtttgtttgctcatctcgttt
ttagcgttaatttctatgaatataagaattttgtctggtttactttttttaatagctgct
atagtcacaattccgcttgtattgttgattgtaactctaatcattcctatcttattcttt
atcattgcaatgatgatgtttgctagaagagatagaatagaaacagtgccatcttattat
aatgaatatgatcaaccatactatgatgagagaggtttttatgaaccagagtcaagaaat
25   gaacatggatataatgatgatgtgtatgaacctatgcatactaaaaaggaagatagaaat
acaagacgtcaattcaatagaaatgctcagcaacaagattcctataatggtataactgat
aatcaacccgatgaagatacatcttccgatcaactttattcagacgaatatgtagataat
gaagataaatattctcaatttccaaaaagagcagttgaaagtgaatatgcatctcaacaa
actgaagatgaaccaacagtcatgtcaagacaagctaagtacaataaaaaatctaaaaat
30   acggattttgaagatgcgcaacaggaacatatggaaggtaatcaatttgatgacatagga
gttgttgaaccacaaattgatcctaaagaactaaaagcgcaaagaaaaagagaaaaagca
gaaatacgtgctaagaaaaaagaaaagagaaaagcatataataaacgtatgaaagaacga
agaaaaaaccagccaagtgctgttaaccaacgacgtatgaattatgaagaacgtcgacaa
atgattaataatgaacaagaagatacagataataacttaaatcaacaggaagattcaaaa
35   aaagaaaattaa

Sequence 2426
MLGEQYTQIKRPANRLTEKILGWFSWVFLLILTIVSMFIALVSFSNDTSIANLENTLNNN
ELVQQILANNDLSTTQFVIWLQNGVWAIIVYFIVCLLISFLALISMNIRILSGLLFLIAA
40   IVTIPLVILIVTLIIPILFFIIAMMMFARRDRIETVPSYYNEYDQPYYDERGFYEPESRN
EHGYNDDVYEPMHTKKEDRNTRRQFNRNAQQQDSYNGITDNQPDEDTSSDQLYSDEYVDN
EDKYSQFPKRAVESEYASQQTEDEPTVMSRQAKYNKKSKNTDFEDAQQEHMEGNQFDDIG
VVEPQIDPKELKAQRKREKAEIRAKKKEKRKAYNKRMKERRKNQPSAVNQRRMNYEERRQ
MINNEQEDTDNNLNQQEDSKKEN*
45

Sequence 2427
Contig_0755_pos_1075_314,
is similar to (with p-value 2.0e-65)
>sp:sp|Q45499|SUHB_BACSU EXTRAGENIC SUPPRESSOR PROTEIN SUHB
50   HOMOLOG. >gp:gp|AF012285|AF012285_41 Bacillus subtilis mobA-
nprE gene region. NID: g3282109. >gp:gp|Z99111|BSUB0008_139
Bacillus subtilis complete genome (section 8 of 21): from 13
94791 to 1603020. NID: g2633699.
atgatgcaagaagagttagacattaaaactaaatcgaacccaaatgatttagttacaaat
55   gtggataaggcgacagagaattatctatatgaaacgattcttcataattatccagatcat
caggttattggcgaagagggacatggtcataatctcgagtatttaaaggggggttatttgg
gttattgatccaattgatggaacacttaattttgttcaccaaaaagaaaattttgccatc
tctattggtatttatcatgatgggaagccttatgcaggttttgtttatgatgtcatgaaa
gatgtttatatcatgcaaaggttggacagggtgcatttgaaaatacacataaaacttgaa

595

```
atgattcaaaatactgaacttaaaagaagtattataggtattaatcccaattggctgacg
aaaccaatactcagtgatatttttagttcaatagtgaatgaggcaagaagtgcacgagca
tatggtagtgcagcattagaaattataagtgtagcgaaggtcaattggcggcttaccta
acacctagactacaaccgtgggattttgcaggtggattgttgattttgaacgaagtaggt
gggataggaaccaacttattaggcgataaaattagacttcaatcaaccgaattcaatatta
atagcaaatcctagccttcatcgtgaaatattaaatcatcatttaaatcagcaaagagat
acgcttattacacttcatgaaaaaaggtttggaaagagatag
```

Sequence 2428
```
MMQEELDIKTKSNPNDLVTNVDKATENYLYETILHNYPDHQVIGEEGHGHNLEYLKGVIW
VIDPIDGTLNFVHQKENFAISIGIYHDGKPYAGFVYDVMKDVLYHAKVGQGAFENTHKLE
MIQNTELKRSIIGINPNWLTKPILSDIFSSIVNEARSARAYGSAALEIISVAKGQLAAYL
TPRLQPWDFAGGLLILNEVGGIGTNLLGDKLDFNQPNSILIANPSLHREILNHHLNQQRD
TLITLHEKRFGKR*
```

Sequence 2429
Contig_0756_pos_1371_1877,
putative peptide of unknown function
```
atgaaagacaacaaacctaataattcgaaattaattcaaacatatttaagtaagaaaact
ttaagatatggtacagcaagtgcattaacattggcactctatttatttaacagtaacgta
actgtgtatgcggatgaaaatactgcaaccaaaatcaaggaacatcaccaaaaacttca
cagacagcacctacaaataatactgaaaatacagatgccacagccataacaacagatcaa
aataataatgatgaagaagaatacgatgcgtcatatgaacttccaattctttatgtaact
gtctggctagatgatcaaggaaatattattaaagatgctgtggaagatgctaaaacccct
gcttcagaaaggcaaccggtgaaaattcctgggtaccaacattatagaacttctgtgagt
gacggaattactaagtttatttatcgtaaaattagcactgcacaatcacctatagttgaa
aatcaaaaacgtatggtcgtggcatag
```

Sequence 2430
```
MKDNKPNNSKLIQTYLSKKTLRYGTASALTLALYLFNSNVTVYADENTANQNQGTSPKTS
QTAPTNNTENTDATAITTDQNNNDEEEYDASYELPILYVTVWLDDQGNIIKDAVEDAKTP
ASERQPVKIPGYQHYRTSVSDGITKFIYRKISTAQSPIVENQKRMVVA*
```

Sequence 2431
Contig_0756_pos_4998_4510,
putative peptide of unknown function
```
atgccaaaagtacatcaagttaaggaaagatttgtgaaattaggggaccaacagtttaaa
gcatttgaaattagatacgatacatacattcattacgtgttgatgtgtgatggtgtagat
ttagcaatgaaacagcgcgtggaagattttgtcagtgcgcaaacatggcatcaacaattt
aaaacgattggcgtcatgcttttttcaacaagataaacaattcatatatccactgatacat
atacctaaaatagatagcttaatctgggaaaatagctgtggttcaggagcggcttctatc
ggtgtgttagttaattatctaacagatcatgatattcaagattacctagttaaccaaccc
ggaggcagtattattgtctcatccagaaagtctggacaaaatgaataccaaacaacgatt
aagtgtcaagtttcaactgtcgcaacaggacaagcatatatagaacaggagacaatgacg
caaatatga
```

Sequence 2432
```
MPKVHQVKERFVKLGDQQFKAFEIRYDTYIHYVLMCDGVDLAMKQRVEDFVSAQTWHQQF
KTIGVMLFQQDKQFIYPLIHIPKIDSLIWENSCGSGAASIGVLVNYLTDHDIQDYLVNQP
GGSIIVSSRKSGQNEYQTTIKCQVSTVATGQAYIEQETMTQI*
```

Sequence 2433
Contig_0756_pos_3690_2413,
putative peptide of unknown function
```
atggttggtagtggaccggtcgctattcaacttgctcgactatgtcatttacatggagaa
catatagttgatatggtgagtcgcgttcatgcatcaaccaaatctaagagagtctttgat
gcttatcaacgtgacggcttttttttcagtaatgactcaaaatgatgcacatcagtgtttt
tcaggtaagtttacggttagacattttttttaaagatgttaaagatattactgaatattat
gacgtggtgattttagcatgtactgccgatgcgtatcgaccgatattacagcaattatct
```

596

aagtccacattaaagcgtattaagcaaatcatcttggtctcaccaacattaggatcacat
atgcttgttaagcaattactatcagatgttcaatgtgaaggtgaagtgatttcattttcc
acttatctaggcgatacccgaatatttgataaagcacaaccacattgtgtcctaaccaca
cgagttaaatcaaaattattcgtaggttcgactcaatctcagtctatgacgttgtgtaag

5 cttaagtctttatttgactatttgaatatagaattaacaacgatggacacaccactacat
gcggagatacataatagttcactttatgtacacccaccattgtttatgaatcaattttca
ttaaaggcggtatttgaagggacgaaagtaccagtatatgtatataagctatttccagag
ggtccaatcacaatgaccttaatacacgaaatgcgattaatgtggcaagaaatgatgatg
atattaaaaaaattaaaggtaccttcggtcaatcttctaaagtttatggtgaaagaaaac

10 taccctatacgttatgagaccatgcgcgaagtagatattgaaaactttaaaaatttacca
gctattcatcaagagtatctactttatgtgcgatatacagcaattttaatcgatccgttt
tctaatccggacgatcaaggtgcatattttgatttttctgccgtaccatacaaacatgtt
gatactgatgaacaaggagtcatacatataccacgcatgccgagtgaagattattatcgt
actttgataattcaagcgattggaagagcattaaacgttgcaacaccgatgattgacaca

15 ttgttattacgttatgaaaatactgttaaacaatactgtgacacacatttacatcaacaa
ctatcaaggcaattcgaattacatcattttaaacaggatttagcgttagtgacgaactac
ttaacttttttataaataa

Sequence 2434
20 MVGSGPVAIQLARLCHLHGEHIVDMVSRVHASTKSKRVFDAYQRDGFFSVMTQNDAHQCF
SGKFTVRHFFKDVKDITEYYDVVILACTADAYRPILQQLSKSTLKRIKQIILVSPTLGSH
MLVKQLLSDVQCEGEVISFSTYLGDTRIFDKAQPHCVLTTRVKSKLFVGSTQSQSMTLCK
LKSLFDYLNIELTTMDTPLHAEIHNSSLYVHPPLFMNQFSLKAVFEGTKVPVYVYKLFPE
GPITMTLIHEMRLMWQEMMMILKKLKVPSVNLLKFMVKENYPIRYETMREVDIENFKNLP

25 AIHQEYLLYVRYTAILIDPFSNPDDQGAYFDFSAVPYKHVDTDEQGVIHIPRMPSEDYYR
TLIIQAIGRALNVATPMIDTLLLRYENTVKQYCDTHLHQQLSRQFELHHFKQDLALVTNY
LTFYK*

Sequence 2435
30 Contig_0756_pos_2362_1841,
is similar to (with p-value 3.0e-63)
>gp:gp|AF076683|AF076683_1 Staphylococcus aureus oligopeptid
e transporter putative substrate binding domain (opp-1A), ol
igopeptide transporter putative membrane permease domain (op
35 p-1B), oligopeptide transporter putative membrane permease d
omain (opp-1C), oligopeptide transporter putative ATPase dom
ain (opp-1D), and oligopeptide transporter putative ATPase d
omain (opp-1F) genes, complete cds; and unknown gene. NID: g
3800817.
40 atgaataaactcacaaaactaagtacagtcatttttgtatctggaattattttagccggt
tgtggaaataacaaagaactaacagagaaaaaagagaataaagtattatcatatacaact
gtcasagatattggagatatgaatccccatgtttatggaggttcaatgtcagcagaaagt
atgatttatgagccgttagttcgcaataccaaggatggtattaagccattattagcaaaa
aaatgggacatttcacctgatggtaagacatatacgtttcatttaagggatgatgtatct

45 tttcatgatggtacgaaatttgatgcagatgcagtgaagaaaaacatcgatgcagtacaa
caaaataagaaactacattcatggttaagactttcaacactgattgatgatgtcaaagtt
aaggataagtatacgatacaactacatttgaaggaagcttatcaacctgcgttagcagaa
ctagctatgccacgaccatacgtttttgattttcaactatag

50 Sequence 2436
MNKLTKLSTVIFVSGIILAGCGNNKELTEKKENKVLSYTTVKDIGDMNPHVYGGSMSAES
MIYEPLVRNTKDGIKPLLAKKWDISPDGKTYTFHLRDDVSFHDGTKFDADAVKKNIDAVQ
QNKKLHSWLRLSTLIDDVKVKDKYTIQLHLKEAYQPALAELAMPRPYVFDFQL*

55 Sequence 2437
Contig_0757_pos_801_1223,
putative peptide of unknown function
atgagaaaatggttaaccttactattaattacaacattggtgttaactgcatgtggtaaa
agtaacgaaaaagcttctttagaaaaaagcattgatcagttgaaaaaagaaaataaggat

ttaaaaaaacagaagaaaaagttacaagagcaaaaggataagcttaaacacaaacaggat
agtctccaagaagatgtaaatgacttgcctgctaaaagcacatcccgagataagaaaaat
aaagataatcatgatgcaaaagaaaagtcttcagataatcaatcgacatctgctaatcat
gatgatcaaactaacaaaataaaaagcaatcaagatgaacatgacagtcaatcctctaaa
5   ccacatacacagcagaagccctcacagaatgatagaaaaataatcatcgacaagaacga
tag

Sequence 2438
MRKWLTLLLITTLVLTACGKSNEKASLEKSIDQLKKENKDLKKQKKKLQEQKDKLKHKQD
10 SLQEDVNDLPAKSTSRDKKNKDNHDAKEKSSDNQSTSANHDDQTNKIKSNQDEHDSQSSK
PHTQQKPSQNDRKNNHRQER*

Sequence 2439
Contig_0757_pos_4581_4994,
15 is similar to (with p-value 5.0e-18)
>gp:gp|AF012906|AF012906_6 Bacillus subtilis yojP gene, part
ial cds; yojQ/S, yojR, yojT, yojU, yojV, yojW, yojX, yojY, y
ojZ, and yokA genes, complete cds. NID: g2522404. >gp:gp|Z99
114|BSUB0011_163 Bacillus subtilis complete genome (section
20 11 of 21): from 2000171 to 2207900. NID: g2634230. >gp:gp|AF
020713|AF020713_166 Bacteriophage SPBc2 complete genome. NID
: g3025478.
atgataggtacttatcaaagtgataaaaactttgaaatgatgaagacttttaagcattgg
attcagactaatcattattggaaatatgttgagaaatacggtgtgttaggtatagcatta
25 gataatcctctccacgttcaaagtaatcaatgtagatatgacgttgtttttgagaatagat
gaaacagtaaatgatcagacaatatctaaaagagattttacaggtggcatatatgctgtg
tttaaagttagtcatacaaaaataaatatagagaagttctttagcaatttagaaaatatt
ttaaatgaaagtcatttgcgtatgagaaatgaaccaattatagagagatacattgaagaa
gagggaacagataaagtgtgtgaaatgttagtgcctatctatgaagtaaattaa
30

Sequence 2440
MIGTYQSDKNFEMMKTFKHWIQTNHYWKYVEKYGVLGIALDNPLHVQSNQCRYDVVLRID
ETVNDQTISKRDFTGGIYAVFKVSHTKINIEKFFSNLENILNESHLRMRNEPIIERVIEE
EGTDFVCEMLVPIYEVN*
35

Sequence 2441
Contig_0757_pos_6065_5280,
is similar to (with p-value 6.0e-63)
>gp:gp|U87792|BSU87792_1 Bacillus subtilis tRNA-Ala, phospha
40 tidylglycerophosphate synthase (pgsA) and CinA (cinA) genes,
complete cds, and RecA (recA) gene, partial cds. NID: g1842
434.
atgatacttgtcgatgatatgtggttaaagtcaactaattttctcggttctcaatcagca
ttcacatttaaagttgttatacagttaggttcagtatttgcggccgcctgggtatttaga
45 gaacgcttcttagaaattttacatattggccaacataaacctgaaccttccacttcggga
gacagacgttcaaaaccacgacgtctgaatttaatacatgtattagtaggtatggtccca
gcagggatttttaggattttttatttgatgatttaattgaaaaatacttatttagtgtacca
acagtcttaattggtttatttataggtgccatttatatgattatagctgataagtattct
aaaactgttcagcatcctcaaacagtagatcaaattaattatttccaagcatttgtcatt
50 ggtatctctcaagcaatagctatgtggcctggatttagtagatccggttcaacgatttca
acaggtgttcttatgaaattgaatcataaagctgcatctgatttcacttttattatgtcg
gtaccaattatgttagctgcaagtggattatctttactaaaacattatgagtatattcat
ttagcacacataccattctacattttaggattttttagcggcatttattgttggattuatt
gcaattaaaacattcttacacttaattaataaagttaagttagtacctttttgctacttat
55 agaattgtcttagttatttttatagcaatcctatacttcggattcggtattggcaaagga
atttaa

Sequence 2442
MILVDDMWLKSTNFLGSQSAFTFKVVIQLGSVFAAAWVFRERFLEILHIGQHKPEPSTSG

DRRSKPRRLNLIHVLVGMVPAGILGFLFDDLIEKYLFSVPTVLIGLFIGAIYMIIADKYS
KTVQHPQTVDQINYFQAFVIGISQAIAMWPGFSRSGSTISTGVLMKLNHKAASDFTFIMS
VPIMLAASGLSLLKHYEYIHLAHIPFYILGFLAAFIVGLIAIKTFLHLINKVKLVPFAIY
RIVLVIFIAILYFGFGIGKGI*

Sequence 2443
Contig_0757_pos_3653_3204,
putative peptide of unknown function
atgaagactaacgtaattatagatggagatgcttgtcctgttgttaattctgtcattgaa
ttgacgaaagggacaggcatttttgttacaattttaagaagttttagccatttttcacaa
caaatacaacccgaacatgtaaaaattgtatacgttgatgacggtcccgatgcggtagac
tataaaatagtcgaacttgctagcaataatgatatcgtcatcacacaagattatggactt
gctagtctactgatagacaaagtacatactgtcatgcatcataaaggaaatatttatcac
tcaaacaacatccaaagcttattaaatcaaagatatctaaatgctcaaataagaagacga
ggtggtcgtcacaaaggccctcctcccttcacaacagaggatagacttaaattcgagcat
gctttttagaaaaatcattaatcaaatatag

Sequence 2444
MKTNVIIDGDACPVVNSVIELTKGTGIFVTILRSFSHFSQQIQPEHVKIVYVDDGPDAVD
YKIVELASNNDIVITQDYGLASLLIDKVHTVMHHKGNIYHSNNIQSLLNQRYLNAQIRRR
GGRHKGPPPFTTEDRLKFEHAFRKIINQI*

Sequence 2445
Contig_0758_pos_2562_3488,
is similar to (with p-value 0.0e+00)
>gp:gp|AF072726|AF072726_1 Staphylococcus aureus putative he
me A synthase (ctaA) gene, complete cds. NID: g3320605.
atgatggggtgtttatcattgtttagaaagcaaaaccttaaatggttaggtgtttttagct
acgattattatgacctttgtacaattaggtggcgccctcgtaactaaaacgggatcagaa
gatggttgtggctcgtcttggcctttatgtaatggcgctttacttccagaaaatttacca
atacaaacaattatagaactgagtcatcgcgcagtatcagccatttcacttatagttgta
ttatggcttgtaattacagcttggaaaaacattggatatattaaagaaatcaaaccactc
tctattattagtgttggttttttattagttcaagcacttgtaggtgctgctgctgtgata
tggcaacaaaatccttatgtattagcgctacattttggtatttcacttatcagtttctct
tctgttttcttaatgacattaattattttctcaattgacaaaaaatatgaagctgacatt
ttatttattcacaaacctttacgtatcttaacttggttaatggctatcatcgtatactta
actatttatacaggtgctttagttagacatactaaatcaagtcttgcttatggtgcttgg
cctattccatttgatgatatcgttcctcataatgcgcatgattgggtacaattttcgcac
agaggtatggcgctcatcacttttatctggattatgattacatttatacacgctattaag
aattattcagataatcgaactgtacgttatggttatactgcatcatttatacttgttatc
cttcaagttattacaggtgctttatcagtcataactaatgtcaatttaattattgcgtta
ttccatgctttgtttatcacttacttattcggaatgattgcttattttattttactaatg
ttaagaacgacgagaagtcaaaaataa

Sequence 2446
MMGCLSLFRKQNLKWLGVLATIIMTFVQLGGALVTKTGSEDGCGSSWPLCNGALLPENLP
IQTIIELSHRAVSAISLIVVLWLVITAWKNIGYIKEIKPLSIISVGFLLVQALVGAAAVI
WQQNPYVLALHFGISLISFSSVFLMTLIIFSIDKKYEADILFIHKPLRILTWLMAIIVYL
TIYTGALVRHTKSSLAYGAWPIPFDDIVPHNAHDWVQFSHRGMALITFIWIMITFIHAIK
NYSDNRTVRYGYTASFILVILQVITGALSVITNVNLIIALFHALFITYLFGMIAYFILLM
LRTTRSQK*

Sequence 2447
Contig_0758_pos_7092_3637,
is similar to (with p-value 0.0e+00)
>gp:gp|D83706|D83706_1 Bacillus stearothermophilus DNA for p
yruvate carboxylase, complete cds. NID: g1695685.
gtgtcttggctttttgaaacaaataaagaaattacttgttgctaaccgtggtgaaatcgcc
attagaattttttagagcggcagcagaattaaatatcagtacagtagcaatttattctaat

gaagataaaagttcgttacatagatataaaagcagatgaatcctatctagttggaagtgat
ttaggacctgctgaaagttatttgaatatcgaacgtatcatcgaagtagctcttcgcgca
ggtgtcgatgcaattcatcctgggtatggttttttaagtgaaaatgaacaatttgcacgc
cgatgtgctgaggaaggcattaaatttataggtccgcatcttgaacatctagacatgttt
5   ggagataaggttaaggctagaacaactgctattaacgctaacttacctgtaatcccgggt
acagatggtcctattgaaagttttgaagctgcagaacagtttgctaatgaagcaggttac
ccacttatgattaaggccacaagcggtggcggtggtaaaggtatgcgaatcgttcgtgaa
tcaagcgaattagaagacgctttccatcgtgcgaaatcagaagccgaaaagtcatttggt
aatagcgaagtttatatcgaaagatatattgataatccaaagcatatagaggttcaagtt
10 attggtgatgaattcgggaatatcattcatttgtatgaaagagattgctccgtacaacga
cgtcatcaaaaggttgttgaagttgcaccttcagtaggtctttctaacaaattaagagag
cgaatttgtgatgccgcaattcaactgatggaaaatataaaatacgtcaacgctggaaca
gtagaatttttagtttctggggatgaattttttcttcattgaggttaatccacgtgttcaa
gttgagcatacaattactgaaatgattactggtatagacattgtgaaaacgcaaatttta
15 gttgctgatggagaatcgttatttggagataaaaatctctatgccacagcaaaatgaaatt
caaacattagggtatgcgatacaatgtcgtataacaactgaagatcctactaatga t ttt
atgccagattctggcacaattattgcatatcgatcaagtggcggtttttggtgtga gactt
gatgcaggggatggattccaaggtgcagaaatttcaccttactacgattcactattagtt
aagctttctacacatgccgtttcatttaaacaagctgaagagaaaatggaacgttcatta
20 cgcgaaatgcgaattcgtggcgtaaagacgaatattccatttctcatcaatgttatgcgt
aatgataaatttagaagtggtgattatactactaaatttattgaagaaacacctgaactt
ttcgatattgcaccgacattggacagaggtaccaagactttagagtatattggtaatgtg
acgataaacggatttcctaatgtagaaaagcgtccaaaaccagaatatgaatctaccaaa
atcccaaaaatttctcaaaagaaaatcaatcagttatttggaacaaaacaaattcttgag
25 caacatggaccaacaggtgttacaaattgggttagagaacaagaagatgttttaattacc
gatactacatttagagatgcacaccaatctttacttgcaacacgtgtaagaacaaaagat
atgatgaacattgcatctaaaactgctgaagttttttaaagatagttttttcattagaaatg
tggggtggtgcaacatttgatgtcgcctataatttcttgaaagagaatccatgggaacgt
ttagaaagattgcgcaaagccattccgaatgtgttattccaaatgttattacgagcttcg
30 aacgcagtaggttataaaaaactatcctgataatgtaattaagaaattcgttcatgaaagt
gcaaaagctggtgtagatgtttttccgtatattcgactcattgaactgggttgatcaaatg
aaagtagcgaatgaagctgttcaagaagctggaattggtatctgagggtacaatttgctat
acaggtgatattttaaatgctgaacgttccaatatttatactttagattattacgttaaa
atggctaaagaactggaaagagaagggattccatatattagcaattaaagatatggctggt
35 ttattgaaaccgaaagcagcttacgaattaattggtgaattacgtgaggcaacacatctt
ccaattcatttacatacacatgatactagtggaaatggattgttgacatataaacaagca
attgatcctggcgtagatattatagatactgctgttgcatctatgagtggtttaacgagt
caaccaagtgcaaattcattatattatgcactaaatggatttccacgtaatttaagaact
gatattgatgggttagaagagttgagtcattactggtctgtagtcagaccttactatgca
40 gactttgagagtgatatcaaatcaccaaatacagaaatttatcaacatgaaatgccaggt
ggccaatattcaaacttaagtcaacaagctaaaagtttaggattgggcgaacgttttgat
gaagtcaaagagatgtatcgtcgtgtcaacttcctgtttggagatcttgtaaaagtaaca
ccatcttcaaaggtagttggagatatggcactatatatggtgcaaaatgatcttgatgaa
gatacggtcatcaatgatggttataaattagatttcccagaatctgttgtgtcattcttt
45 aaaggtgacattggacaacctgtcaacggattcaacaagaaattgcaagatgttatttta
aaaggacagcaaccaattactgaaagaccaggtgaatacttggagccggtcgatttttgaa
gcaatccgtcaagaattaagcgacatacaacaagacgaggtaacagaacaagatataatt
agttatgtactttatccgaaggtatataaacaatatattcaaacgaaagagcaatttggt
aatgtatctttactggatacaccgacattcttatttggcatgcgtaatggtgaaacagtt
50 gaaattgaaattgatactggtaaacgtctaattattaaattagaaacaatcagtgaacca
gatgagaatggtaaacgtacaatttattacgctatgaatggtcaagcaagacgtatttat
attcaagatgaaaatgttaaaacgaatgctaatgttaaacctaaggcggataaatcaaat
ccaaatcatattggtgctcaaatgcctggttctgtaactgaagtcaaagtgtctgtaggc
gatgaagttcaagctaatcagccattattaatcactgaagcaatgaagatggaaacgacg
55 attcaggcaccatttgatggaattattaaacaaatcaatgttgctaatggagatgccatt
gcca agyagatttattagtggaaattgaaaagtaa

Sequence 2448
VSWLLKQIKKLLVANRGEIAIRIFRAAAELNISTVAIYSNEDKSSLHRYKADESYLVGSD

600

LGPAESYLNIERIIEVALRAGVDAIHPGYGFLSENEQFARRCAEEGIKFIGPHLEHLDMF
GDKVKARTTAINANLPVIPGTDGPIESFEAAEQFANEAGYPLMIKATSGGGGKGMRIVRE
SSELEDAFHRAKSEAEKSFGNSEVYIERYIDNPKHIEVQVIGDEFGNIIHLYERDCSVQR
RHQKVVEVAPSVGLSNKLRERICDAAIQLMENIKYVNAGTVEFLVSGDEFFFIEVNPRVQ

5   VEHTITEMITGIDIVKTQILVADGESLFGDKISMPQQNEIQTLGYAIQCRITTEDPTNDF
MPDSGTIIAYRSSGGFGVRLDAGDGFQGAEISPYYDSLLVKLSTHAVSFKQAEEKMERSL
REMRIRGVKTNIPFLINVMRNDKFRSGDYTTKFIEETPELFDIAPTLDRGTKTLEYIGNV
TINGFPNVEKRPKPEYESTKIPKISQKKINQLFGTKQILEQHGPTGVTNWVREQENVLIT
DTTFRDAHQSLLATRVRTKDMMNIASKTAEVFKDSFSLEMWGGATFDVAYNFLKENPWER

10   LERLRKAIPNVLFQMLLRASNAVGYKNYPDNVIKKFVHESAKAGVDVFRIFDSLNWVDQM
KVANEAVQEAGMVSEGTICYTGDILNAERSNIYTLDYYVKMAKELEREGFHILAIKDMAG
LLKPKAAYELIGELREATHLPIHLHTHDTSGNGLLTYKQAIDAGVDIIDTAVASMSGLTS
QPSANSLYYALNGFPRNLRTDIDGLEELSHYWSVVRPYYADFESDIKSPNTEIYQHEMPG
GQYSNLSQQAKSLGLGERFDEVKEMYRRVNFLFGDLVKVTPSSKVVGDMALYMVQNDLDE

15   DTVINDGYKLDFPESVVSFFKGDIGQPVNGFNKKLQDVILKGQQPITERPGEYLEPVDFE
AIRQELSDIQQDEVTEQDIISYVLYPKVYKQYIQTKEQFGNVSLLDTPTFLFGMRNGETV
EIEIDTGKRLIIKLETISEPDENGKRTIYYAMNGQARRIYIQDENVKTNANVKPKADKSN
PNHIGAQMPGSVTEVKVSVGDEVQANQPLLITEAMKMETTIQAPFDGIIKQINVANGDAI
ATGDLLVEIEK*

20

Sequence 2449
Contig_0758_pos_2303_1368,
is similar to (with p-value 3.0e-59)
>sp:sp|P24009|COXX_BACSU PROBABLE CYTOCHROME C OXIDASE ASSEM

25   BLY FACTOR. >gp:gp|Z98682|BS16823KB_3 Bacillus subtilis geno
mic DNA 23.9kB fragment. NID: g2339988. >gp:gp|X54140|BSCTAB
F_1 B. subtilis ctaB-F genes for cytochrome a assembly facto
r and cytochrome-c oxidase (EC 1.9.3.1) subunits II, I, II,
and IVB. NID: g994793. >gp:gp|Z99111|BSUB0008_160 Bacillus s

30   ubtilis complete genome (section 8 of 21): from 1394791 to 1
603020. NID: g2633699.
atgagaaatttaaggagggggaattatgaacaaagatcaaactttgtcacatactacgggc
cgtgtatccttcaaagaattacaacaaattattaaaatgggccttgttcaaggtaattta
atacctgcttttgcaggcgcatggcttgcaatagtaatgacaaaccattccttcctatct

35   tccattccacaaatactattgatgctagttggctctacgcttattatggggggcgcttgt
gctttaaataattattatgatcaagatattgatcgcattatgcctagtaagcaaagtaga
ccaacagtaaatgatagaatatctgatagaaacttattaatgttaagttttgggatgatg
ttaataggtgaagcatgtttattcttattaaatataccttctggtgttttaggattaatt
ggtattgttggatatgtatcttactattcaatttggtctaagcgccatacaacttggaat

40   actgttgttggaagtttcctggagctgtaccaccattaattggttgggtagctatcgat
ggatcattaagtttagcagcagtagcactcttttttagttgtctttgttggcaacctatc
catttctacgctctagcaattaaacgtagtgatgagtatgcgcttgcaaatattcctatg
ttaccatcagtgaaaggtttcaaacggacaagagtaagcatgtttatttggttagtgtta
ttattaccattgccattcttattatctaatttaggcgtaactttgttgttattgctaca

45   ctacttaatttaggatggttagctttaggttttacaacgttcagaaaagaatctaatcaa
actaaatgggcaacgcaaatgttcgtttattcattgaactacttagtagtattctttgca
ctcgttgtagttgtttcattaatcaagatgatataa

Sequence 2450
50   MRNLRRGIMNKDQTLSHTTGRVSFKELQQIIKMGLVQGNLIPAFAGAWLAIVMTNHSFLS
SIPQILLMLVGSTLIMGGACALNNYYDQDIDRIMPSKQSRPTVNDRISDRNLLMLSFGMM
LIGEACLFLLNIPSGVLGLIGIVGYVSYYSIWSKRHTTWNTVVGSFPGAVPPLIGWVAID
GSLSLAAVALFLVVFCWQPIHFYALAIKRSDEYALANIPMLPSVKGFKRTRVSMFIWLVL
LLPLPFLLSNLGVTFVVIATLLNLGWLALGFTTFRKESNQTKWATQMFVYSLNYLVVFFA

55   LVVVVSLIKMI*

Sequence 2451
Contig_0758_pos_1351_881,
putative peptide of unknown function

```
atgaaacttatgaacgtacccattttaccaacgataagtacatcgtgtattgttattagt
gcgattttagtcgctattggttgggcattgatttggaaacgtcaagttcataagcataaa
aacattatgctatgggctgccttttcgctttaacattctttattatttatgcagcaaga
actattttttatcggtaatacagctttcggtggaccaagttctattaaagtttattacact
```
5
```
attttcttagtttttccatatcattcttgcaacagttggtggcgtttttaggcttaattcaa
atcattttagccttcaaagataaacttcatattcacagaaaaattgggccttgggcttca
ataatttggttctttaccgcaattactggtgttgcagtttatgtattgttatatgtattg
tatccaggtggagaaacaacatcattgcttaaagctacattaggtctataa
```

10   Sequence 2452
MKLMNVPILPTISTSCIVISAILVAIGWALIWKRQVHKHKNIMLWAAFFALTFFIIYAAR
TIFIGNTAFGGPSSIKVYYTIFLVFHIILATVGGVLGLIQIILAFKDKLHIHRKIGPWAS
IIWFFTAITGVAVYVLLYVLYPGGETTSLLKATLGL*

15   Sequence 2453
Contig_0759_pos_5038_3959,
is similar to (with p-value 2.0e-29)
>sp:sp|P33642|YFIT_PSEAE HYPOTHETICAL 39.5 KD OXIDOREDUCTASE
   IN FIMT 3'REGION (DADA*) (ORFZ). >gp:gp|L48934|PSEPILRV_2 P
20   seudomonas aeruginosa (isolate pRIC351) pilR gene, 3' end of
   cds, dada*, fimT, fimU and pilV genes, complete cds. NID: g
   1161217.
```
atgtcaattgctagacacctcagtgcaacacacttagatgttgcagtcatagatagagat
gtacctggaaagcatgcgtcatataaagctggaggtatgcttggcgcacaaaatgaattt
```
25
```
acagaggatagtgacttgtttcaattagccatcgaatctcgtgctatgtttccacaatta
agtaaatcattattagatgaaacaggcatagacattcaatttaaaaattcaggacttatc
aaaattgctaatgaacacgatgatatctcatctataaaacgacaatatcaatttctgaat
agtcaagaccgtagtgtcaaacaattatcagatgatgatttgctacaacttacacatggt
gaagttaaaccttcatacgcggccattcacataccacacgatggtcaaattaatgcacat
```
30
```
cattacacactggcattattagaatcaatgaagttaagagatattaagcgttatgagtct
acagaggtcacttcaatagaacggcataatggctattattcagtgaaaaccgatcaatct
tcaacaattgaagcgcacaaaattatcgttgcaggtggcgcatggtcttcgcaattatta
acacaatatcatctacaacgacaagtgattggcgttaaaggtgaagttatcttattagaa
aataacgatctttcacttactgagacattatttatgactaatggttgttacatcgttcca
```
35
```
aaacaacccaatcgtttttaattggtgcgacgagtgaatttaataattattctgtcggt
actacagatgaaggtatggattggcttcttcgccatgcatatcatcgtgtacctcaacta
aaagacagtcatatactgaagaaatggtcaggagtaagaccatacacagaaaaagaaatg
ccagtcatggatcaaattgatgatggcttatacgtgataagtggtcattatcgaaacgga
atattattgtcacctattatcggtcgtgacattgccaattggctactttctggtattaaa
```
40
```
ccatcacgttattcaagttttacagttacaaggaggaataatcatgaagtgtatcattaa
```

Sequence 2454
MSIARHLSATHLDVAVIDRDVPGKHASYKAGGMLGAQNEFTEDSDLFQLAIESRAMFPQL
45   SKSLLDETGIDIQFKNSGLIKIANEHDDISSIKRQYQFLNSQDRSVKQLSDDDLLQLTHG
EVKPSYAAIHIPHDGQINAHHYTLALLESMKLRDIKRYESTEVTSIERHNGYYSVKTDQS
STIEAHKIIVAGGAWSSQLLTQYHLQRQVIGVKGEVILLENNDLSLTETLFMTNGCYIVP
KQPNRFLIGATSEFNNYSVGTTDEGMDWLLRHAYHRVPQLKDSHILKKWSGVRPYTEKEM
PVMDQIDDGLYVISGHYRNGILLSPIIGRDIANWLLSGIKPSRYSSFTVTRRNNHEVYH*
50

Sequence 2455
Contig_0759_pos_3004_2006,
is similar to (with p-value 2.0e-52)
55   >gp:gp|AF012285|AF012285_3 Bacillus subtilis mobA-nprE gene
   region. NID: g3282109. >gp:gp|Z99111|BSUB0008_99 Bacillus su
   btilis complete genome (section 8 of 21): from 1394791 to 16
   03020. NID: g2633699.
```
atgtcgcgttatgaacgtcaaacacgctttgcaccatttggagaagagggtcagcaaaag
```

ctatcctcctctcaaatacttattttggcgctggtgctttaggaagccatattgtagat
caactcgcacgcatgggggctcatcatattgcaatcgtcgatatggatattgttgaaatt
tcaaatttacatcgacaaacactcttcgatgaagaagacgcacatactttaatatccaaa
gttgaagcaatcaagcataaggttaatcaaattaatataaatgtcaatctaacaacttat

5 gatttagaagttacttcatcaaatatcgaaaatttgataaaaaatgtcgaaccagacatc
atcattgatggcatggataacttcaaaatacgatacctgattaatgaggtttgtcacaag
tatcaaatcccatgggtttatggtgcagctgttggtagtaaaggatcagtatatggaata
gatcaccaaggaccatgtctaaaatgtttattgcaaacaattcctgacacaggggaaagt
tgcgctattaatggcgtaattccccctgttatatcaatgattgcaagctatgaagtagca

10 gaggccgtacgttatctttcaggaaaaggattttcaaagcaattaatcactattgatgca
tttaatatcaattataagtcaatgaatgtagatgcactcaaaaataaagattgcccagtg
tgtgaaaaacatgaatatacgttactagaaagccaacaagaacgtactattgaggacttg
tgtgggaatgcttatttatttagattcccacctaaagctttaaacacgctgcccatttc
cctgggaatatggtgaaatctacttcctttgccaaattaattcaatatcaaacttatgaa

15 ttcaccttgtttaaagatggtcgtatgaatgcatatggtatacacaatgatgaagaagca
catcacctatacaatacgttgttaaaatccatacgctaa

Sequence 2456
MSRYERQTRFAPFGEEGQQKLSSSQILIFGAGALGSHIVDQLARMGAHHIAIVDMDIVEI

20 SNLHRQTLFDEEDAHTLISKVEAIKHKVNQININVNLTTYDLEVTSSNIENLIKNVEPDI
IIDGMDNFKIRYLINEVCHKYQIPWVYGAAVGSKGSVYGIDHQGPCLKCLLQTIPDTGES
CAINGVIPPVISMIASYEVAEAVRYLSGKGFSKQLITIDAFNINYKSMNVDALKNKDCPV
CEKHEYTLLESQQERTIEDLCGNAYLFRFPPKAFKHAAHFPGNMVKSTSFAKLIQYQTYE
FTLFKDGRMNAYGIHNDEEAHHLYNTLLKSIR*

25

Sequence 2457
Contig_0759_pos_1898_1284,
is similar to (with p-value 4.0e-38)
>sp:sp|Q48630|APL_LACLA ALKALINE PHOSPHATASE LIKE PROTEIN. >

30 pir:pir|S39339|S39339 alkaline phosphatase-like protein - La
ctococcus lactis >gp:gp|Z29065|LLALPHLP_2 L.lactis (MG1363)
apl gene for alkaline phosphatase like protein. NID: g435295

atggaacaaattatcactgattttattagtaagtggggttatacagcgatattcatttta

35 atcttattagagaacgtattacctgtcgttccatctgagattattttaacttttgcaggc
ttattatctgtgaaatcacacttatctatttggacattattaatcatagcaacaattgct
tcattcattggtttactcatttttgtattatatttgtagacttatctcagaagagaaatta
tatcgtttcgttgatcgacatggtaagtggatgaagttaaaaagtaaagatttgaaacgg
gcaaatgattggtttaaaaagtatggtgcgtgggctgtattttatgtcgtttttgtccca

40 gtacttcgagtattaattacaatacctgctggcattaatcgaatgaacgttatacagttt
acaactttatctttaataggtactacaatttggaattttgctttaatactgctcggtcgt
ttgctcagtgacagttttgacgctttgatgaatggtattcatacatattcacgtatcatg
tatgtcattattattattgcagtcatatattttgttatacgttatttaatgaaacgtcgt
cggagtgttaaataa

45

Sequence 2458
MEQIITDFISKWGYTAIFILILLLENVLPVVPSEIILTFAGLLSVKSHLSIWTLLIIATIA
SFIGLLILYYICRLISEEKLYRFVDRHGKWMKLKSKDLKRANDWFKKYGAWAVFLCRFVP
VLRVLITIPAGINRMNVIQFTTLSLIGTTIWNFALILLGRLLSDSFDALMNGIHTYSRIM

50 YVIIIIAVIYFVIRYLMKRRRSVK*

Sequence 2459
Contig_0759_pos_1086_151,
is similar to (with p-value 0.0e+00)

55 >gp:gp|D78193|BACGNTZA_11 Bacillus subtilis 36kb sequence be
tween gntZ and trnY genes encoding 34 ORFs. NID: g1064780. >
gp:gp|Z99124|BSUB0021_124 Bacillus subtilis complete genome
(section 21 of 21): from 3999281 to 4214814. NID: g2636442.
atgaaacaaaaatatctagatttactctcacaaaaatttgacagtgcagaaaaacttgct

```
actgaaattattaacttagagtcaatcttagaattacctaaagggactgaacattttgtt
ag'.gacc.tcatggtgaatacgaatctttccaacatgtttttaagaaacggatctg\/aaat
gtgcgtgctaaaattaatgatatcttcaaagataaattatcccagcaagaaatca\cgac
ttagcagcattagtatactatccggaagaaaaactaaaattagttaaaaataatttcgat
```
5
```
tcaatcggaacattaaatatttggtatattacaaccattcaacgattaattgatttaatt
acatattgctcatcaaaatatacacgttcaaaattacgcaaagcattacctgaacaatac
gtttatattattgaagagctactttacaagagcaatgaatttcataataaaaagccttat
tatgaaacattagttaaccaaattattgaattagaacaatcagatgatttaatcattggc
ctttcctatactgtacaacgtctagtcgtagaccatcttcatgtcgtgggcgatatctat
```
10
```
gaccgtggtcctaaacctgataagattatggatacattaataaattatcattctgtagat
atccaatgggggaaatcatgatgtattatggattggcgcctatgctggttcaaaagtatgt
cttgctaaccttctacgtatctgtgcacgttatgataatttagatattattgaagatgca
tatggcatcaatctacgccctttacttacgcttgctgaaaagtattacgatgctgaaaac
ccagcgtttaaacctaagaaacgaccagataaagacgtcagtcttacaaaacgcgagaaa
```
15
```
gtcaaatcacaaaaattcatcaagcaattgcgatga
```

Sequence 2460
MKQKYLDLLSQKFDSAEKLATEIINLESILELPKGTEHFVSDLHGEYESFQHVLRNGSGN
VRAKINDIFKDKLSQQEINDLAALVYYPEEKLKLVKNNFDSIGTLNIWYITTIQRLIDLI
20
TYCSSKYTRSKLRKALPEQYVYIIEELLYKSNEFHNKKPYYETLVNQIIELEQSDDLIIG
LSYTVQRLVVDHLHVVGDIYDRGPKPDKIMDTLINYHSVDIQWGNHDVLWIGAYAGSKVC
LANLLRICARYDNLDIIEDAYGINLRPLLTLAEKYYDAENPAFKPKKRPDKDVSLTKREK
VKSQKFIKQLR*

25
Sequence 2461
Contig_0762_pos_5747_4368,
is similar to (with p-value 0.0e+00)
>sp:sp|P25811|THDF_BACSU POSSIBLE THIOPHENE AND FURAN OXIDAT
ION PROTEIN THDF. >pir:pir|JQ1215|JQ1215 hypothetical 50K pr
30
otein - Bacillus subtilis >gp:gp|D26185|BAC180K_60 B. subtil
is DNA, 180 kilobase region of replication origin. NID: g467
326. >gp:gp|X62539|BSORIGS_5 B.subtilis genes rpmH, rnpA, 50
kd, gidA and gidB. NID: g40020. >gp:gp|Z99124|BSUB0021_207 B
acillus subtilis complete genome (section 21 of 21): from 39
35
99281 to 4214814. NID: g2636442.
```
atggatttgatacgattacaagtatttcaacaccgatgggtgaaggtgctattggaatt
gtgagattatctgggccacaagctattgaaatcggagatatcttatataaaggtaagaaa
aagttatctgaagttgagacgcatacaataaaattacggtcatattattgatccagaaaca
aatgaaacagttgaagaagtcatggtgtctgtattacgtgcccctaaaactttcacacga
```
40
```
gaagatattattgagataaaattgtcatggtggtattttaacaattaatcgtatattagag
tt\aactatgacttatggtgcacgtatggcagaaccaggtgaatatacaaaacgtg\attt
ttaaatggtcgtatagatttatctcaagcagaagcggttatggatttttatacgtt\:caa
actgatcgagcttctaaggttgcgatgaatcaaatagaaggacgtttaagtgacttaatc
aagaaacaacgtcaatccatattagagatactcgcccaagttgaagttaacattgattat
```
45
```
ccagagtatgatgatgtagaagacgcaacgacggacttcttactagaacagtctaagcgt
attaaagaagaa\tcaatcagttacttgaaacaggagcacaaggtaaaataatgagagaa
gggttatctacagttattgtaggacgtcctaatgttgggagtcttcgatgctaaataac
cttattcaagataataaagcaattgtgactgaggtcgctggtacaacaagagacgtgtta
gaagaatatgtcaatgttagaggtgtcccgttacgacttgtagatactgcgggtattagg
```
50
```
gatactgaagatatcgtagagaagattggtgtagaacgttctaggaaagctttaagtgaa
gcagatttaattttatttgtgcttaataacaatgaacctctgacggaagatgatcaaact
ttattcgaagtcattaaaaatgaggatgttattgtaatcattaataaaacagatttagaa
cagcgattagatgttagcgaactaagagagatgattggtgatatgccacttatacaaaca
tcgatgcttaaacaagaaggtattgatgaattagaaatacaaattaaagatttattcttt
```
55
```
ggtggcgaagtacaaatcaagatatgacttatgtatctaattcacgtcacatttcattg
ttgaaacaagcgagacaatcaattcaagatgcgattgatgctgctgagtctggtatccca
atggatatggtacagattgatttaacacgtacttgggaaattctaggagaaattattgga
gaatcagcgagtgatgaattaatagatcaactatttagtcaattttgtttaggaaaataa
```
```

Sequence 2462
MDFDTITSISTPMGEGAIGIVRLSGPQAIEIGDILYKGKKKLSEVETHTINYGHIIDPET
NETVEEVMVSVLRAPKTFTREDIIEINCHGGILTINRILELTMTYGARMAEPGEYTKRAF
LNGRIDLSQAEAVMDFIRSKTDRASKVAMNQIEGRLSDLIKKQRQSILEILAQVEVNIDY
PEYDDVEDATTDFLLEQSKRIKEEINQLLETGAQGKIMREGLSTVIVGRPNVGKSSMLNN
LIQDNKAIVTEVAGTTRDVLEEYVNVRGVPLRLVDTAGIRDTEDIVEKIGVERSRKALSE
ADLILFVLNNNEPLTEDDQTLFEVIKNEDVIVIINKTDLEQRLDVSELREMIGDMPLIQT
SMLKQEGIDELEIQIKDLFFGGEVQNQDMTYVSNSRHISLLKQARQSIQDAIDAAESGIP
MDMVQIDLTRTWEILGEIIGESASDELIDQLFSQFCLGK*

Sequence 2463
Contig_0762_pos_4295_2418,
is similar to (with p-value 0.0e+00)
>sp:sp|P25812|GIDA_BACSU GLUCOSE INHIBITED DIVISION PROTEIN
A. >pir:pir|JQ1216|BWBSGA gidA protein - Bacillus subtilis >
gp:gp|D26185|BAC180K_59 B. subtilis DNA, 180 kilobase region
 of replication origin. NID: g467326. >gp:gp|X62539|BSORIGS_
 6 B.subtilis genes rpmH, rnpA, 50kd, gidA and gidB. NID: g40
020. >gp:gp|Z99124|BSUB0021_206 Bacillus subtilis complete g
enome (section 21 of 21): from 3999281 to 4214814. NID: g263
6442.
gtggttcaagaatatgatgtagtagtcattggtgctggtcacgccggtattgaagcaggt
ctagcttcagctcgccggggtgctaaaacactgatgttaacaattaatttagataatatt
gctttcatgccatgtaatccatctgtaggtggtcctgcgaaaggaatcgttgtacgtgaa
atagacgctttaggtggacaaatggcaaaaactattgataaaactcacattcaaatgcgt
atgcttaatacaggtaaaggtccagctgttagagctttacgtgctcaagcagataaagta
ttatatcaacaagaaatgaagcgtgtacttgagaatgaggataatttagacatcatgcaa
ggtatggttgatgaactcattatagaagataatgaagttaaaggtgttcgtactaatatt
ggtacagaatatcgttctaaagctgtcattattacaacaggtacattcttacgtggagaa
attatactaggaaacttaaaatattctagtggccctaaccatcaattaccatctgtaact
ctagcggataatttaagaaaattaggatttgatatcgttagatttaaaacgggtacacca
cccagtcgaaatgcgagaaccatcgattattctaaaactgaaatccaaccaggtgacgat
ataggtcgagcgtttagttttgaaacaaccgaatttattttagatcaattaccttgttgg
ttaacttatacaaatggagatacacatcaagtcattgatgataacttacatttatctgct
atgtattccggtatgattaaaggtacaggtcctagatattgtccatcaattgaggataaa
tttgtccgctttaacgataaaccaagacatcaactttttcttagaacctgaaggacgtaat
acgaatgaggtatacgtgcaaggattatctactagtttacctgaacatgttcaacgtcaa
atgttagaaaccattccaggtcttgaaaaagcagatatgatgcgtgcgggttatgctatt
gaatatgatgcaatcgtgcctactcaattatggccaacgttagaaacaaaagcgattaaa
aacttgtatactgcaggtcagattaatggaacatcaggatatgaagaagcagcgggacaa
ggaatcatggcaggtattaacgctgctggtaatgttttaggtacaggtgaaaaaatactc
agccgttcagacgcatatattggtgtacttatagatgatttagtcactaagggtacaaat
gaaccgtatcgattattaacttcacgtgcggaatatcgattattactacgtcatgataat
gctgatttacgtcttactgatatggggtatgaattaggttttaatatcagaagaacgctat
gcaagatttaatgaaaagcgtcaacaaatcaaagatgaaatacaacgacttaccgatgta
cgtattaaaccaaatgaacatacgcaagcaattattgaagctaagggtggttcaagatta
aaagatggcatattagcgattgatttattacgtcgtcccgaaatgaactacgaaacaatt
ttagaaatcttagaagaatcacatcaacttcctgaagcggttgaggaacaagttgaaatt
caaacaaaatatgaaggttatatcaataaatctttacaacaagtagaaaaagttaaaaga
atggaagcgaaaaaaattcctgaggatttagattatagcaaggtagatagtttagcatct
gaagcacgcgaaaagttagctgaagttaaaccattaaatattgcacaggcttcacgaatt
tcagtcgctgaatccagcagatatctcaattctacttgtttatttagaacaaggtaaactt
caaagggtgaaacaataa

Sequence 2464
VVQEYDVVVIGAGHAGIEAGLASARRGAKTLMLTINLDNIAFMPCNPSVGGPAKGIVVRE
IDALGGQMAKTIDKTHIQMRMLNTGKGPAVRALRAQADKVLYQQEMKRVLENEDNLDIMQ
GMVDELIIEDNEVKGVRTNIGTEYRSKAVIITTGTFLRGEIILGNLKYSSGPNHQLPSVT

LADNLRKLGFDIVRFKTGTPPRVNARTIDYSKTEIQPGDDIGRAFSFETTEFILDQLPCW
LTYTNGDTHQVIDDNLHLSAMYSGMIKGTGPRYCPSIEDKFVRFNDKPRHQLFLEPEGRN
TNEVYVQGLSTSLPEHVQRQMLETIPGLEKADMMRAGYAIEYDAIVPTQLWPTLETKAIK
NLYTAGQINGTSGYEEAAGQGIMAGINAAGNVLGTGEKILSRSDAYIGVLIDDLVTKGTN
5   EPYRLLTSRAEYRLLLRHDNADLRLTDMGYELGLISEERYARFNEKRQQIKDEIQRLTDV
RIKPNEHTQAIIEAKGGSRLKDGILAIDLLRRPEMNYETILEILEESHQLPEAVEEQVEI
QTKYEGYINKSLQQVEKVKRMEAKKIPEDLDYSKVDSLASEAREKLAEVKPLNIAQASRI
SGVNPADISILLVYLEQGKLQRVKQ*

10   Sequence 2465
      Contig_0762_pos_1424_816,
      is similar to (with p-value 1.0e-40)
      >sp:sp|P37524|YYAA_BACSU HYPOTHETICAL 32.8 KD PROTEIN IN SPO
      0J-GIDB INTERGENIC REGION. >pir:pir|S18078|S18078 hypothetic
15   al protein 3 - Bacillus subtilis >gp:gp|D26185|BAC180K_57 B.
       subtilis DNA, 180 kilobase region of replication origin. NI
      D: g467326. >gp:gp|X62539|BSORIGS_8 B.subtilis genes rpmH, r
      npA, 50kd, gidA and gidB. NID: g40020. >gp:gp|Z99124|BSUB002
      1_204 Bacillus subtilis complete genome (section 21 of 21):
20   from 3999281 to 4214814. NID: g2636442.
      atgttcgaaattatagccggcgaacgacgatttcgagcattacagtcgttgcataaacct
      caagtagatgtcattgttcgagatatggatgatgaagaaacagcggtagttgcattgatt
      gaaaatattcaacgtgaaaacttatctgtcgtcgaagaagcggaagcttataaaaagtta
      cttgaaatcggggggaacgactcaaaatgaattagcaaaaagtctaggcaagagccaaagc
25   ttcattgctaataaacttagattattgaagttagcacccaatgtgattaagagatnacgt
      gaaggtaagattacagaaagacatgcacgagcggtattagtattacctgatgaaacacaa
      gaagaattaatcgagcaagttattagtcagaagttgaatgtgaaacaaactgaggataga
      gtacgtcagaaaactggaccagagaaagtaaaagcgcagactttccaattttctcaagat
      gtaacacaagcaaaagaagaactaggtaagagtattgaaacgatagaaaaatcaggtata
30   cgcgttgaacaaaaagataaagaacatgaagattattatgaaattaaaataaagatatat
      aagaaataa

      Sequence 2466
      MFEIIAGERRFRALQSLHKPQVDVIVRDMDDEETAVVALIENIQRENLSVVEEAEAYKKL
35   LEIGGTTQNELAKSLGKSQSFIANKLRLLKLAPNVIKRLREGKITERHARAVLVLPDETQ
      EELIEQVISQKLNVKQTEDRVRQKTGPEKVKAQTFQFSQDVTQAKEELGKSIETIEKSGI
      RVEQKDKEHEDYYEIKIKIYKK*

      Sequence 2467
40   Contig_0763_pos_624_1304,
      is similar to (with p-value 2.0e-26)
      >gp:gp|U76260|PAU76260_1 Peptostreptococcus asaccharolyticus
       alpha- and beta-subunits of L-serine dehydratase (sdhB) and
       (sdhA) genes, complete cds. NID: g2315864.
45   atggctaaaagctatgattatcaaagtgctttcgatattattggaccagtaatgatggga
      ccttcaagttctcatacagcaggtgcagtaaaaattggtaattcagcgagagctgtgtta
      ggagatatgcctaagcatatagaaattcgttattatgaatcttttgctaaaacgcatcaa
      gggcatggtacagacgttgctattgtcggaggtgctatgggctacagcactttcgatagt
      agaattaaatcatccttagacatagcaaaagatgaaaatattacaattgatattattgaa
50   gatgaaggagaaagtattggtcaacatcctaactgtgcttatatcaaagcaaatacgaaa
      gacggacgttatatagaagtgataggtatttctattggtggcggtacaatcaaactaaaa
      ggtatcaatgtaaatggtttaaatgtggaactgaatcatgggcttccaatgttdgttata
      gatggaaatatgaataaagctaaaataaatcatcttattaatgatttatcagatatggac
      ttagacttaggtgaagaattaatagaaacaaatgataatgaaggtttagttgtatttcct
55   ttaaataaagcaatctcagaatcagcattaaatattattaaagataaacatagtgattta
      aacgtttcctatatcaaatag

      Sequence 2468
      MAKSYDYQSAFDIIGPVMMGPSSSHTAGAVKIGNSARAVLGDMPKHIEIRYYESFAKTHQ

GHGTDVAIVGGAMGYSTFDSRIKSSLDIAKDENITIDIIEDEGESIGQHPNCAYIKANTK
DGRYIEVIGISIGGGTIKLKGINVNGLNVELNHGLPMLVIDGNMNKAKINHLINDLSDMD
LDLGEELIETNDNEGLVVFPLNKAISESALNIIKDKHSDLNVSYIK*

Sequence 2469
Contig_0763_pos_3838_4449,
putative peptide of unknown function
atgaagcagtggatgaatagattaatcaccttaataggcgtattgttaatcattttagct
atttatttattctcaaagccatatatcgataattatctacatgaaaaagataacgatcat
aaaattgaaaattatgataaaaaggaaaaagaacagacaaagacatctaaatcgacgcca
aagataccttccgataaatctaaaatggctggttatatagaagttccagatgcacaaata
aaagaaccagtataccctggtccagcaacaccagaacaactcaatagaggtgttagtttt
gcagaaggtgacgaatctcttaatcaacagaatatttcaattgctggtcatacgtttaca
gatcgttcgcactatcaatttacaaattaaaatcagccaaatcggtagtaaagtgtat
tttaaaactggaaatcaaactagaaagtataaaataactaaaatacgtgatgttaagcct
acagaggttaaggtattagacgaacatcctaataagaaaaatcaattaacattaattact
tgcgatgactataacgaagaaacgggtgtttgggaaacaaggaaaatattcatagctaca
caaattaactaa

Sequence 2470
MKQWMNRLITLIGVLLIILAIYLFSKPYIDNYLHEKDNDHKIENYDKKEKEQTKTSKSTP
KIPSDKSKMAGYIEVPDAQIKEPVYPGPATPEQLNRGVSFAEGDESLNQQNISIAGHTFT
DRSHYQFTNLKSAKIGSKVYFKTGNQTRKYKITKIRDVKPTEVKVLDEHPNKKNQLTLIT
CDDYNEETGVWETRKIFIATQIN*

Sequence 2471
Contig_0763_pos_4794_5603,
putative peptide of unknown function
atgattaaagccattgcggtagatatggatggaacatttcttgacacaaataaacagttt
gatcgaaatcgttttgaaactattttaaagaattaatagataaaaatattaagtttata
gctgcgagtgggaatcaatttgcaaagctaaatcaatttttggagatagggaaatgttc
tttatatctgaaaatggagcagtcatctataaaggtaatcaactttacaattatcgaagt
tttgatcagtatatttttcaaaaagttgtaaattatttaaatttgaatcaaaagataaac
aatttgattatttgtggtgtaaaaagtgcatatattttaaaagaaacaagcgaagcattt
aagcaagatgcacgtacatattatcaccaactaatagaggttgactccttacaaacatta
cctgatgatgattatgtgaaaattgctttcaatataaatcgtcagactcatccagactta
gatgagaaattagctcttaagtttaaagacgatattaaactagtatcaagtgggagagat
agtatagatgttattatgccaaatatgactaagggtcaagctttgtctagattattaaaa
gaatggcaaatgcctgcttcacatttaatggcatttggagatgcaaataacgataaagat
atgttggagcttgccgaacatagttatgttatggctaatagtgaagatcaatcattattt
aatatagcgagtcatgtggcaccttccaatgatgaacaaggcgtactatcaacaatcgaa
aatgttgttctcggttattccaataaataa

Sequence 2472
MIKAIAVDMDGTFLDTNKQFDRNRFETIFKELIDKNIKFIAASGNQFAKLKSIFGDREMF
FISENGAVIYKGNQLYNYRSFDQYIFQKVVNYLNLNQKINNLIICGVKSAYILKETSEAF
KQDARTYYHQLIEVDSLQTLPDDDYVKIAFNINRQTHPDLDEKLALKFKDDIKLVSSGRD
SIDVIMPNMTKGQALSRLLKEWQMPASHLMAFGDANNDKDMLELAEHSYVMANSEDQSLF
NIASHVAPSNDEQGVLSTIENVVLGYSNK*

Sequence 2473
Contig_0763_pos_7023_7691,
is similar to (with p-value 2.0e-20)
>gp:gp|AF012552|AF012552_2 Helicobacter pylori prolipoprotei
n diacylglycerol transferase (lgt) and NADPH-linked flavin n
itroreductase (rdxA) genes, complete cds. NID: g2564440.
atgattatgaatcagatgaatcaaacgattattgatgcattccattttagacatgcgaca
aaagaatttgaccctacgaaaaaaattagtgatgaagattttaatacgatttttagaaaca
ggtagattatctccaagttcactaggtttagaaccttggcactttgtagtggttcaaaat

```
      aaagaattgagagaaaaattgaaagcctatagttgggggagcacaaaagcaacttgataca
      gcaagtcactttgtattaattttttgctcgtaagaatgtgacggctcatacagattacgtg
      caacatttacttcgtggcgtcaaaaaatatgaagaaagtacaattccagcagttgaaaat
      aaatttgatgatttccaagaaagtttccatattgccgataatgaacgaacattatatgac
5     tgggcgagtaaacaaacatatattgcattagcaaacatgatgacaagtgctgcattacta
      ggtatcgactcatgtccaattgaaggatttgatttagataaagtgactgaaattctttca
      gatgagggtgttttagatacggaacaatttggtatttcagttatggtaggctttggttac
      agagcacaagaacctaaacatggcaaagttagacaaaacgaagacgacatcattagttgg
      attgaataa
10

      Sequence 2474
      MIMNQMNQTIIDAFHFRHATKEFDPTKKISDEDFNTILETGRLSPSSLGLEPWHFVVVQN
      KELREKLKAYSWGAQKQLDTASHFVLIFARKNVTAHTDYVQHLLRGVKKYEESTIPAVEN
      KFDDFQESFHIADNERTLYDWASKQTYIALANMMTSAALLGIDSCPIEGFDLDKVTEILS
15    DEGVLDTEQFGISVMVGFGYRAQEPKHGKVRQNEDDIISWIE*

      Sequence 2475
      Contig_0763_pos_6789_5797,
      is similar to (with p-value 0.0e+00)
20    >gp:gp|U31175|SAU31175_1 Staphylococcus aureus D-specific D-
      2-hydroxyacid dehydrogenase (ddh) gene, complete cds. NID: g
      1644432.
      atgacaaaaattatgttttcggcacaagagcatatgagaaggacatggcattacgttgg
      ggaaagaaaaataatatcgatgtcactacatcaacagaacttttaagtgtagatactgtc
25    gatcaattaaaagattatgacggtgttacaacaatgcagttcggtaaattagaacctgaa
      gtttaccctaaattagagtcctatggtattaaacaaattgcacaacgtacggctggattt
      gatatgtatgacttagaacttgcaaaaaaacatgaaattattatctcgaatataccaagt
      tattcacctgaaacaattgctgaatattcggtatctatcgctctgcaactcgtacgaaaa
      ttcccaacaattgaaaaacgtgtgcaagcacataatttcacatgggcgtcccctattatg
30    tctcgtccagtaaaaaatatgactgtagcaatcatcggtacagggcgtattggtgctgca
      actggtaaaatctatgctggttttggtgcgagagtagttggttatgatgcatatcctaat
      cattctttatctttcttagaatataaagaaacagtagaggatgcaattaaagatgctgat
      attatctcattacatgtacccgctaataaagatagtttccatttatttgataacaatatg
      tttaaaaatgttaaaaaaggtgccgtttttagtcaatgccgcaagaggagctgtgataaac
35    acgcctgatttaattgaagcagtaaataatggtacattatcaggtgctgccattgacaca
      tatgaaaatgaagctaattatttcacatttgattgttcaaatcaaacgattgacgaccca
      atattattagacctaattagaaatgaaaatatttagttacacctcatattgccttttc
      tccgatgaagcagtacaaaattagtagagggtggtttgaatgcagcattatcagtaatt
      aatactggcacatgtgatacgcgattaaactaa
40

      Sequence 2476
      MTKIMFFGTRAYEKDMALRWGKKNNIDVTTSTELLSVDTVDQLKDYDGVTTMQFGKLEPE
      VYPKLESYGIKQIAQRTAGFDMYDLELAKKHEIIISNIPSYSPETIAEYSVSIALQLVRK
      FPTIEKRVQAHNFTWASPIMSRPVKNMTVAIIGTGRIGAATGKIYAGFGARVVGYDAYPN
45    HSLSFLEYKETVEDAIKDADIISLHVPANKDSFHLFDNNMFKNVKKGAVLVNAARGAVIN
      TPDLIEAVNNGTLSGAAIDTYENEANYFTFDCSNQTIDDPILLDLIRNENILVTPHIAFF
      SDEAVQNLVEGGLNAALSVINTGTCDTRLN*

      Sequence 2477
50    Contig_0763_pos_3669_3175,
      is similar to (with p-value 3.0e-17)
      >pir:pir|JT0409|JT0409 phosphinothricin-N-acetyltransferase
      - Streptomyces viridochromogenes >gp:gp|M22827|STMPAT_2 Stre
      ptomyces viridochromogenes phosphinothricin N-acetyltransfer
55    ase (pat) gene, complete cds. NID: g295177. >gp:gp|X65195|SV
      PTT_3 S.viridochromogenes genes pms, phsA, pat and dea for p
      hosphinomethylmalic-acid-synthase, phosphinothricin-tripepti
      de-synthetase A, phosphinothricin-N-acetyltransferase and N-
      acetylphosphinothricin-tripeptide deacetylase, respectively.
```

NID: g47997. >gp:gp|A02774|A02774_1 Artificial phosphinothr
icin resistance gene. NID: g345279. >gp:gp|A02804|A02804_1 S
.viridochromogenes phosphinothricin resistance gene. NID: g3
45154. >gp:gp|A29201|A29201_1 Synthetic DNA for phosphinothr
5   icin resistance gene (viral/herbicide resistance) from paten
t WO9111517. NID: g1248925.
atgattagatttgcacgactagaagatcttcaagatattttgacaatttataatgatgcc
atccttaatacaacagctgtttatacgtataagccacaacaattagatgaacgtcttcaa
tggtatcaatctaaagcaaaaataaacgaacctatatgggtttatgaaaaagaagggaaa
10   gtagttggttttgccacttatggttcctttagacaatggccggcctatttatatactatt
gaacattctatatatgttcatcaacagtacagaggactaggtatcgcttctcaattatta
gagaatttaattcgttacgctaaagaacaaggttatcgcaccattgttgctgggattgat
gcatcgaacatggatagtatcgcattgcataagaagtttgacttctcacatgcaggtaca
attaaaaatgtaggttataaatttgatcgatggctcgatttatcattttatcaatatgat
15   ttatctgattcataa

Sequence 2478
MIRFARLEDLQDILTIYNDAILNTTAVYTYKPQQLDERLQWYQSKAKINEPIWVYEKEGK
VVGFATYGSFRQWPAYLYTIEHSIYVHQQYRGLGIASQLLENLIRYAKEQGYRTIVAGID
20   ASNMDSIALHKKFDFSHAGTIKNVGYKFDRWLDLSFYQYDLSDS*

Sequence 2479
Contig_0764_pos_4801_5163,
is similar to (with p-value 2.0e-52)
25   >sp:sp|Q05217|SECY_STACA PREPROTEIN TRANSLOCASE SECY SUBUNIT
. >pir:pir|S30115|S30115 secY protein - Staphylococcus carno
sus >gp:gp|X70086|SCSECY_1 S.carnosus secY gene. NID: g49188
.
atgattatttatgtagttttaattattgcatttgcatattttatgcttttgtacaagtt
30   aatcctgaaaaaatggcagataaccttaaaaagcaaggtagttatgtcccaggaattaga
cctggtgaacaaacaaaaaaatatattactaaagtactttatagattgacttttgttggt
tcaattttcttagcagctatagctattttacctataattgcgactaaatttatgggctta
ccacaatcaattcaaattggtggtacgagtcttttgatcgttattggtgtagctattgaa
actatgaaaactttagaagcacaagtcactcaaaaagaatataaaggctttggtggtaga
35   taa

Sequence 2480
MILYVVLIIAFAYFYAFVQVNPEKMADNLKKQGSYVPGIRPGEQTKKYITKVLYRLTFVG
SIFLAATAILPIIATKFMGLPQSIQIGGTSLLIVIGVAIETMKTLEAQVTQKEYK.FGGR
40   *

Sequence 2481
Contig_0764_pos_6903_5197,
is similar to (with p-value 2.0e-26)
45   >sp:sp|P54159|YPBR_BACSU HYPOTHETICAL 137.4 KD PROTEIN IN BC
SA-DEGR INTERGENIC REGION. >gp:gp|L77246|BACYACA_6 Bacillus
subtilis (YAC10-9 clone) DNA region between the serA and kdg
loci. NID: g1256615. >gp:gp|Z99115|BSUB0012_144 Bacillus su
btilis complete genome (section 12 of 21): from 2195541 to 2
50   409220. NID: g2634478.
atgaaacgtattaatgaagttggtataccaattattttgtgattaatcagattgataaa
cataatgaagaagaaattacatttgaaacttttaaatcaagagtcgaaaaatcaatcaaa
gactgggatatcaaacttcaagatacttattacgtttcaaagtttgatcatccacagaat
gaaattgacaaactttcaaattttctagtatttatggatcaacatcgtgaatcaacagaa
55   gactatgttaatagaaccattcaattcattaccgacgcacaatacatatacattcaaaat
gaaatgcaatctattcttgacacccttcaaattaatgaagaacaattcgaggaagcatat
attcaatttcaacaaatcaagaagtcagcgcagaagcacaattgctcaatgactctaat
caattatttaattatttaaaacagaagcgtaaagatatattagataatgcttatacatg
acgtacgatatgcgtgaatctttacggaattatttagaaagcatggcaactgattctaaa

gtgaatggattttttaataaaaggaagaaaaaagaagaagaacaaatcaaacgacttaat
gaggcgaccactcaattgcaagagaaagttaatcaacaagtacgacaaccacttcgtgaa
gatatgtcatttttaactagattcataaataaacatgctgtgaatgaaaaaatactaaat
caagaatatgacgtcgttccgtcacttatatcagagctatatcaaactcaaacgagcatt
5   agcaacacatacgttttaacattttcagatgaagttataaaagctttgaataaaaaaata
gaaaatgagtcaacaccactatttgaagaagctgtcaatcatgtacaagttaatgaatta
tcgagtgatgaaaatgaagataggtatgaatatgatagatacattgaacttaacacatta
aaggattcgcttacatcccacaactacaaacattactatatccatttagacgattcttta
gataaattaattggaagaacagagactcattttgaacttaaacaagaaattcaactgct
10   tatcatcgtaaacatgagacacaacatcgtaacgagtttgttacatctaatcaagatatt
aagcgtgcattagatatcgttaaagatgtaccattatttgatcgcactaaacaagatatc
accgataccattctgagactcgataatcaaataacaaaagttggtgttttggtacattt
agcgcaggtaaaagcagcttaatcaacgctctactaggtgaaaactatttagttagttct
cctaatccaacaacagcggcaacgacggaattatcatatggtaaagagagtcaaatcaca
15   ttaaaatcgaaagaacaattactagaggaagttaatcatgtactagaattttatgaaata
tcgtttaacacattagacgactttattgagagtgatttagataagttaaaattgaaacta
gaaaagaaccaacttgcatttattagtgcaattgagaaacattatgaaatgtacacatct
atgttagaacattcacttatacacacagtatcgcttgaagaaattaaaaaatggagtgcc
gaggatgagtatgctactttcgtgaaaactgtacaccttaagctacctttagattggctc
20   aagggtaaaatcattattgcccattaa

Sequence 2482
MKRINEVGIPIIFVINQIDKHNEEEITFETFKSRVEKSIKDWDIKLQDTYYVSKFDHPQN
EIDKLSNFLVFMDQHRESTEDYVNRTIQFITDAQYIYIQNEMQSILDTLQINEEQFEEAY
25   IQFQQNQEVSAEAQLLNDSNQLFNYLKQKRKDILDNAYIMTYDMRESLRNYLESMATDFK
VNGFFNKRKKKEEEQIKRLNEATTQLQEKVNQQVRQPLREDMSFLTRFINKHAVNEKILN
QEYDVVPSLISELYQTQTSISNTYVLTFSDEVIKALNKKIENESTPLFEEAVNHVQVNEL
SSDENEDRYEYDRYIELNTLKDSLTSHNYKHYYIHLDDSLDKLIGRTETHFELKQENSTA
YHRKHETQHRNEFVTSNQDIKRALDIVKDVPLFDRTKQDITDTILRLDNQITKVGVFGTF
30   SAGKSSLINALLGENYLVSSPNPTTAATTELSYGKESQITLKSKEQLLEEVNHVLEFYEI
SFNTLDDFIESDLDKLKLKLEKNQLAFISAIEKHYEMYTSMLEHSLIHTVSLEEIKKWSA
EDEYATFVKTVHLKLPLDWLKGKIIIAH*

Sequence 2483
35   Contig_0764_pos_4419_4120,
is similar to (with p-value 5.0e-52)
>sp:sp|P43148|SEPA_STAEP EXTRACELLULAR ELASTASE PRECURSOR (E
C 3.4.24.-) (SEPP1). >pir:pir|A40659|A40659 elastase, SepP1=
etrxtracellular metalloprotease - Staphylococcus epidermidis
40   >gp:gp|X69957|SESEPP1A_1 S.epidermis gene for protease. NID
: g396258.
atgtctaatccagagcgttttggacaaccatctcatatgaatgattttgtttatacaaat
tctgacaacggaggcgtacatacgaattcaggtattccgaacaaagcagcttacaacaca
attcgtagtattggtaaacaacgttctgaacaaatttattatagagcgttaactgtttat
45   ttaacttcaaattctgatttccaagatgccaaagcatcattacaacaagcagcatttgat
ttatatggcgacggtattgctcaacaagtaggtcaagcatgggacagtgttggcgtgtag

Sequence 2484
50   MSNPFRFGQPSHMNDFVYTNSDNGGVHTNSGIPNKAAYNTIRSIGKQRSEQIYYRLTVY
LTSNSDFQDAKASLQQAAFDLYGDGIAQQVGQAWDSVGV*

Sequence 2485
Contig_0764_pos_3606_3145,
55   putative peptide of unknown function
atgaaaaagagtaaacgacaagatttagtaactatgattgttaagcaaaatcacatttat
aaaaaagcagatattattgattacattgatgatcactttggtgtacgttatagcatgact
actattgcgagagatttaagagaacttcatatttatcgcctgcctgtgaaggcaaatcaa
tatgaatacaaattacttacgcaacaatctcaattagactcaagagtaagactaaatgat

tatatagaaacagaaattattaacactatgattaaagaatcgtatatactaataaagacg
acaccaggatttgcacaaagcataaattattatattgatcagttacaactgaaagagatt
ataggaacgataagtggtaatgatacaattatgattcttacacattcccagtctatagct
gaatatgtctattacaaaatatttaatcataattattcataa

Sequence 2486
MKKSKRQDLVTMIVKQNHIYKKADIIDYIDDHFGVRYSMTTIARDLRELHIYRLPVKANQ
YEYKLLTQQSQLDSRVRLNDYIETEIINTMIKESYILIKTTPGFAQSINYYIDQLQLKEI
IGTISGNDTIMILTHSQSIAEYVYYKIFNHNYS*

Sequence 2487
Contig_0764_pos_2812_1742,
is similar to (with p-value 0.0e+00)
>gp:gp|Y17554|BLY17554_1 Bacillus licheniformis arcA, arcB,
arcC and arcD genes. NID: g3687415.
atgagtatgacaaatggacctattcaagttaacagtgaaatcggcaaattaaagacggta
ttacttaaaagaccagggaaggaattagaaaacttagtgcctgattacttagatggatta
ttgtttgatgatattccattttttaaaggtagctcaacaagagcatgatcattttgctcaa
gttcttcaagatgaaggaatagaagtgctttatttagaaaaattagcagcacaaagtata
gaagattcaaatgtcagagagcaatttatagatgacgttttagcagaatctagaaaaact
atcctaggtcatgaaaaagaaataaaaaaactcttttcaactttgtcaaatcaagcatta
attaataagattatggctggcgtacgcaaagaagaaatacaacttgaatcgacacatctt
gtagagtacatggatgataaatatccgttttatcttgacccaatgcctaatctatatttc
acacgtgatcctcaagcttcaattggtagaggtatgacagtaaatcgtatgtttggaga
gcgagacgcagagaatcgattttcatttcatatattttaaaacatcatcctagatttaaa
gatgagaatattcctttatgggtggacgtcgtgactgtccgttcaacatcgaaggtggagac
gaactggtgttatctaaagatgtacttgcaataggatatctgaacgtacttctgcacaa
gcaattgaacgtttagcacgacgtattgttttaaagatccgttatctacttttaaaaaggtg
gtggcgattgagattccaactagtcgaacatttatgcacttagatactgtttgtacaatg
attgattacgacaaattcactacacattcagcaattcttaaatcagaaggaaacatgaat
atctttattatcgaatatgatgataaagctgaagatatcaaaatccaacattctagtcat
cttaaacaaacattagaagaagtgctcgatgttgatgaaatcacattaataccaactgga
aatggtgatatcatcgacggtgctcgtgaatggatgtccaggtcctggtga

Sequence 2488
MSMTNGPIQVNSEIGKLKTVLLKRPGKELENLVPDYLDGLLFDDIPFLKVAQQEHDHFAQ
VLQDEGIEVLYLEKLAAQSIEDSNVREQFIDDVLAESRKTILGHEKEIKKLFSTLSNQAL
INKIMAGVRKEEIQLESTHLVEYMDDKYPFYLDPMPNLYFTRDPQASIGRGMTVNRMFWR
ARRRESIFISYILKHHPRFKDENIPLWVDRDCPFNIEGGDELVLSKDVLAIGISERTSAQ
AIERLARRIFKDPLSTFKKVVAIEIPTSRTFMHLDTVCTMIDYDKFTTHSAILKSEGNMN
IFTISYDDKAEDIKIQHSSHLKQTLEEVLDVDEITLIPTGNGDIIDGAREWMSRSH*

Sequence 2489
Contig_0765_pos_772_1692,
putative peptide of unknown function
gtgatttcatctactggttgtttttgttacttttttctgttggttcaccttcgccaactttt
tcccctgttaatgggttcttagttgttggtgttgtaattgttttttgttcctggttcacct
ttctgtttaacgcgctctttacctggtttttaaatcaggattgaattcacgtttcttgtcg
aatggaatttcttccgttgacgtgatcggatctccatcaactggaccatattttgtcaca
tcatccactggtggtgtgactacttcgcctgtatcagggttttttaactcctggtttacct
ggaacgtcctcttggctacctttcggtgcatttggatcaaattcatccttatggcctggc
ttgatttcttcgccaccatattctgtgatttcatctactggttgttttgttatttttttct
gttggttcaccttcgccaacttttttccctgttaatgggttcttagttgttggtgttgta
attgtttttgttcctggttcaccttttttgtttaacacgctcttcacctggttttaaatca
ggattgaattcacgtttcttgtcgaatggaatttcttccgttgacgtgatcggatctcca
tcaactggaccatattttgtcacatcatccacaggtggagtaactacttcgcctgtatca
ggatttttaaccccccggcttacctggttgcgttgtttgactacctttcggtgcatttgg
atcaaattcatccttatggcctggcttgatttcttcgccaccataatgaacgatttcatc
cactggttgttttgttatttttttctgttggttcaccttcgccaactttttctcctgtatt

aggattgacataagttggtgttgttgttgtttcaattcctggttcaccttttttggactac
tttttctgtacctggggctaa

Sequence 2490
5   VISSTGCFVTFSVGSPSPTFSPVNGFLVVGVVIVFVPGSPFCLTRSLPGFKSGLNSRFLS
NGISSVDVIGSPSTGPYFVTSSTGGVTTSPVSGFLTPGLPGTSSWLPFGAFGSNSSLWPG
LISSPPYSVISSTGCFVIFSVGSPSPTFSPVNGFLVVGVVIVFVPGSPFCLTRSSPGFKS
GLNSRFLSNGISSVDVIGSPSTGPYFVTSSTGGVTTSPVSGFLTPRLTWLRCLTTFRCIW
IKFILMAWLDFFATIMNDFIHWLFCYFFCWFTFANFFSCIRIDISWCCCCFNSWFTFLDY
10 FFCTWG*

Sequence 2491
Contig_0765_pos_3052_2720,
putative peptide of unknown function
15 atgcagaagcacctcaatctgagccaacgaagacagaagaaggaagcaacgcaaaagcag
ctcaatctgagccaacgaaggcagaagaaggaggcaatgcagaagcagctcaatctgagc
caacgaagacagaagaaggaagcaatgcagaagcacctcaatctgagccaacgaaggcag
aagaaggaggcaatgcagaagcacctcaatctgagccaacgaagacagaagaaggaggca
atgcagaagcaccgaatgttccaactatcaaagctaattcagataatgatacacaaacac
20 aattttcagaagcccctacaagaaatgacctag

Sequence 2492
MQKHLNLSQRRQKKEATQKQLNLSQRRQKKEAMQKQLNLSQRRQKKEAMQKHLNLSQRRQ
KKEAMQKHLNLSQRRQKKEAMQKHRMFQLSKLIQIMIHKHNFQKPLQEMT*
25
Sequence 2493
Contig_0765_pos_2621_1443,
putative peptide of unknown function
gtgaatttaaattatagttctccgtttatgtccttattaagcatgcctgctgatagttca
30 tccaataacactaaaaatacaatagatataccgccaactacggttaaaggtagagataat
tacgattttttacggtagagtagatatcgaaagtaatcctacagatttaaatgcgacaaat
ttaacgagatataattatggacagccacctggtacaacaacagctggtgcagttcaattt
aaaatcaagttagttttgataaagatttcgactttaacattagagtagcaaacaatcgt
caaagtaatacaactggtgcagatggttggggctttatgttcagcaagaaagatggggat
35 gatttcctaaaaaacggtggtatcttacgtgaaaaaggtacacctagtgcagctggtttc
agaattgatacaggatattataataacgatccattagataaaatacagaaacaagctggt
caaggctatagagggtatgggacatttgttaaaaatgactcccaaggtaatacttctaaa
gtaggatcaggtactccatcaacagattttcttaactacgcagataatactactaatgat
ttagatggtaaattccatggtcaaaaattaaataatgttaatttgaaatataatgcttca
40 aatcaaactttacagctacttatgctggtaaaacttggacggctacgttatctgaatta
ggattgagtccaactgatagttacaatttttttagttacatcaagtcaatatggaaatggt
aatagtggtacatacgcaagtggcgttatgagagctgatttagatggtgcaacattgaca
tacactcctaaagcagtcgatggagatccaattatatcaactaaggaaataccatttaat
aagaaacgtgaatttgatccaaacttagccccaggtacagaaaaagtagtccaaaaaggt
45 gaaccaggaattgaaacaacaacaacaccaacttatgtcaatcctaatacaggagaaaaa
gttggcgaaggtgaaccaacagaaaaaataacaaaacaaccagtggatgaaatcgttcat
tatggtggcgaagaaatcaagccaggccataaggatgaatttgatccaaatgcaccgaaa
ggtagtcaaacaacgcaaccaggtaagccggggggttaa

50 Sequence 2494
VNLNYSSPFMSLLSMPADSSSNNTKNTIDIPPTTVKGRDNYDFYGRVDIESNPTDLNATN
LTRYNYGQPPGTTTAGAVQFKNQVSFDKDFDFNIRVANNRQSNTTGADGWGFMFSKKDGD
DFLKNGCILREKGTPSAAGFRIDTGYYNNDPLDKIQKQAGQGYRGYGTFVKNDSQGNTSK
VGSGTPSTDFLNYADNTTNDLDGKFHGQKLNNVNLKYNASNQTFTATYAGKTWTATLSEL
55 GLSPTDSYNFLVTSSQYGNGNSGTYASGVMRADLDGATLTYTPKAVDGDPIISTKEIPFN
KKREFDPNLAPGTEKVVQKGEPGIETTTTPTYVNPNTGEKVGEGEPTEKITKQPVDEIVH
YGGEEIKPGHKDEFDPNAPKGSQTTQPGKPGG*

Sequence 2495

Contig_0765_pos_0_1408,
putative peptide of unknown function
gtggatgatgtgacaaaatatggtccagttgatggagatccgatcacgtcaacggaagaa
attccattcgacaagaaacgtgaattcaatcctgatttaaaaccaggtgaagagcgtgtt

5   aaacaaaaaggtgaaccaggaacaaaaacaattacaacaccaacaactaagaacccatta
acaggggaaaaagttggcgaaggtgaaccaacagaaaaaataacaaaacaaccagtagat
gaaatcacagaatatggtggcgaagaaatcaagccaggccataaggatgaatttgatcca
aatgcaccgaaaggtagccaagaggacgttccaggtaaaccaggagttaaaaaccctgat
acaggcgaagtagtcacaccaccagtggatgatgtgacaaaatatggtccagttgatgga

10   gatccgatcacgtcaacggaagaaattccattcgacaagaaacgtgaattcaatcctgat
ttaaaaccaggtaaagagcgcgttaaacagaaaggtgaaccaggaacaaaaacaattaca
acaccaacaactaagaacccattaacaggggaaaaagttggcgaaggtgaaccaacagaa
aaagtaacaaaacaaccagtagatgaaatcacagaatatggtggcgaagaaatcaagcca
ggccataaggatgaatttgatccaaatgcaccgaaaggtagccaagaggacgttccaggt

15   aaaccaggagttaaaaatcctgatacaggcgaagtagttactccaccagtggatgatgtg
acaaaatatggtccagttgatggagatccgattacgtcaacggaagaaattccgtttgat
aaaaaacgcgaatttgatccaaacttagcgccaggtacagagaaagtcgttcaaaaaggt
gaaccaggaacaaaaacaattacaacaccaacaactaagaacccattaacaggggaaaaa
gttggcgaaggtgaaccaacagaaaaagtaacaaaacaaccagtggatgaaatcgttcat

20   tatggtggcgaagaaatcaagccaggccataaggatgaatttgatccaaatgcaccgaaa
ggtagccaagaggacgttccaggtaaaccaggagttaaaaaccctgatacaggcgaagta
gttactccaccagtggatgatgtgacaaaatatggtccagttgatggagatccgattacg
tcaacggaagaaattccgtttgataaaaaacgcgaatttgatccaaacttagcgccaggt
acagagaaagtcgttcaaaaaggtgaaccaggaacaaaaacaattacaacaccaacaact

25   aagaacccattaacaggggaaaaagttggcgaaggtgaaccaacagaaaaagtaacaaaa
caaccagtggatgaaatcgttcaTATCG

Sequence 2496
VDDVTKYGPVDGDPITSTEEIPFDKKREFNPDLKPGEERVKQKGEPGTKTITTPTTKNPL

30   TGEKVGEGEPTEKITKQPVDEITEYGGEEIKPGHKDEFDPNAPKGSQEDVPGKPGVKNPD
TGEVVTPPVDDVTKYGPVDGDPITSTEEIPFDKKREFNPDLKPGKERVKQKGEPGTKTIT
TPTTKNPLTGEKVGEGEPTEKVTKQPVDEITEYGGEEIKPGHKDEFDPNAPKGSQEDVPG
KPGVKNPDTGEVVTPPVDDVTKYGPVDGDPITSTEEIPFDKKREFDPNLAPGTEKVVQKG
EPGTKTITTPTTKNPLTGEKVGEGEPTEKVTKQPVDEIVHYGGEEIKPGHKDEFDPNAPK

35   GSQEDVPGKPGVKNPDTGEVVTPPVDDVTKYGPVDGDPITSTEEIPFDKKREFDPNLAPG
TEKVVQKGEPGTKTITTPTTKNPLTGEKVGEGEPTEKVTKQPVDEIVHIX

Sequence 2497
Contig_0766_pos_1949_3403,
40   is similar to (with p-value 1.0e-62)
>sp:sp|P14567|DP3A_SALTY DNA POLYMERASE III, ALPHA CHAIN (EC
2.7.7.7). >pir:pir|A45915|A45915 DNA-directed DNA polymeras
e (EC 2.7.7.7) III alpha chain - Salmonella typhimurium >gp:
gp|M29701|STYDNAE_1 S.typhimurium polymerase III polymerase
45   subunit gene, complete cds. NID: g153951.
atggaagaaataccaacttatataacccgtagacataatcctaaccaatttgcttattta
catccagatttagaaccaatcttaaaaaacacatatggtgttatcatttatcaagaacaa
ataatgctaatagcaagtcaagttgctggttttagttatggtgaagcagatattttaaga
agggcaatgagtaaaaagaatcgtgcaatcttagaaagtgagcgtcaacatttcattgat

50   ggtgcaaaaaataacggttacgatgaacagataagtaagcaaatttttgatttaatactt
aagtttgcagattatgggttcccacgtgcccatgctgttagttactcaaaaattgcatac
attatgagctatttaaaagtgcactatcctcattattttttatgcaaatatcttgagtaat
gtaataggaagtgaaaaaaagactgcagctatgattgacgaagctaagcaccaaagaatt
agcatcttgcctcccaatattaatcaaagtcattggtattataaggcaagtaataaagga

55   atatatctgtctttaggtacaattaaaggaattggatatcaaagcgttaaattaattatt
gatgaacgtcagcagaatggaccttatagagatttctttgattttttcaagacgtatacca
aaaagggtgaaaaatagaaaattacttgagtctcttatcttagtaggcgcattcgacact
tttggcaaaactagagcgacattattacaagcaattgatcaagtattagatttgaattct
gatgttgagcaagatgaaatgctttttcgatctttttaactcctaaacaatcgtatgaagaa

aaagaggaactacctgatcaattattaagtgattatgaaaaagaatacctaggattctat
attagtaaacatccagttgaaaagaaatttgaaaagaaacaatatttaggcatatttcaa
ttgtctaatggaagtcactaccaacctatacttgttcaatttgaccatatcaaacaaata
agaacgaagaatggtcaaaatatggcatttgtaacgatgaatgatggaagaacgatgatg
5   gatggagtgattttcccagataagtttaaaaaatacgaaacttctatttcaaaggaacag
atgtatatcgtattaggtaaatttgaaaagcgtaaccaacaaatgcaacttatcatcaat
caacttttttgaagttgaagcgtatgagcaaacaaaattgtctaattcgaaaaaagttatt
ttacgtaatgtaacacatctagaaccacaatttgaacattcaaaagtagaatctaatgaa
caacatgcattaaatatttatggttttgacgaaagtgcaaataagatgacaatgttgggga
10   caaattgaacgtcaacgtcaaaattttgatctattaatacaaacttattcgccagctgat
attagatttatttaa

Sequence 2498
MEEIPTYITRRHNPNQFAYLHPDLEPILKNTYGVIIYQEQIMLIASQVAGFSYGEADILR
15   RAMSKKNRAILESERQHFIDGAKNNGYDEQISKQIFDLILKFADYGFPRAHAVSYSKIAY
IMSYLKVHYPHYFYANILSNVIGSEKKTAAMIDEAKHQRISILPPNINQSHWYYKSSNKG
IYLSLGTIKGIGYQSVKLIIDERQQNGPYRDFFDFSRRIPKRVKNRKLLESLILVGAFDT
FGKTRATLLQAIDQVLDLNSDVEQDEMLFDLLTPKQSYEEKEELPDQLLSDYEKEYLGFY
ISKHPVEKKFEKKQYLGIFQLSNGSHYQPILVQFDHIKQIRTKNGQNMAFVTMNDGRTMM
20   DGVIFPDKFKKYETSISKEQMYIVLGKFEKRNQQMQLIINQLFEVEAYEQTKLSNSKKVI
LRNVTHLEPQFEHSKVESNEQHALNIYGFDESANKMTMLGQIERQRQNFDLLIQTYSPAD
IRFI*

Sequence 2499
25   Contig_0766_pos_3567_4364,
is similar to (with p-value 7.0e-79)
>sp:sp|P54537|YQIZ_BACSU PROBABLE AMINO-ACID ABC TRANSPORTER
 ATP-BINDING PROTEIN IN BMRU-ANSR INTERGENIC REGION. >gp:gp|
 D84432|BACJH642_255 Bacillus subtilis DNA, 283 Kb region con
30   taining skin element. NID: g2627063. >gp:gp|Z99116|BSUB0013_
107 Bacillus subtilis complete genome (section 13 of 21): fr
om 2395261 to 2613730. NID: g2634723.
atgattgtactattgttctacatatttactagtaaatcgtgtatatttaaaatgattcag
ttttcttttcagccagtgattaatattaaaaatttaaataaaaaatttggagcaaatgaa
35   gtattgagagatattaatcttactgttgaaaagggtgaagtggttgcaataattggacca
tctggaagtggtaaaagcactttactccgttgtatgaatttgttagatgtaccttcaaaa
ggtaaagttatatttgaagataatgaattaactcaacataatgttcatttagataattta
cgacaaaaaatgggtatggtatttcaaaattttaatttatttcctcataaaaaggtcatt
gaaaatgtaatgttggcaccactttttattacataaagatagtaaagatcaattaaaagag
40   aaagctttatatttacttgaaaaagtggggcttaaagacaaagcagattcatatcctaat
caactgtctggaggtcaaaaacaaagagttgctattgcaagagctttggcaatggaacct
gatgttatgttatttgatgaaccaacatctgcacttgatcctgaagtagtaggggatgtt
ttaaaagtaatgagacaattagcgaatgaaggtatgacaatggtgattgtcacgcatgag
atgaactttgctaaagaaataagtgataaagtagtatttatggccgatggtgttgtggtt
45   gaatctggtacaccacaaaacatatttgaaaatcctcagcacagtcgaactgaaaatttt
ttatcacgagtgttataa

Sequence 2500
MIVLLFYIFTSKSCIFKMIQFSFQPVINIKNLNKKFGANEVLRDINLTVEKGEVVAIIGP
50   SGSGKSTLLRCMNLLDVPSKGKVIFEDNELTQHNVHLDNLRQKMGMVFQNFNLFPHKKVI
ENVMLAPLLLHKDSKDQLKEKALYLLEKVGLKDKADSYPNQLSGGQKQRVAIARALAMEP
DVMLFDEPTSALDPEVVGDVLKVMRQLANEGMTMVIVTHEMNFAKEISDKVVFMADGVVV
ESGTPQNIFENPQHSRTENFLSRVL*

55   Sequence 2501
Contig_0766_pos_4706_5149,
is similar to (with p-value 1.0e-20)
>pir:pir|S05373|S05373 ctc protein - Bacillus subtilis (frag
ment)

```
atggtttcagattatcaatttgatccattaaaaaaccaaatcactcatattgactttta
gcaatcaacatgagtgaagaacgtactgttgaagtacaagttcaattagttggtgaagct
gtaggtgctaaagaaggcggcgtagttgaacaaccattattcaacttagaagttacagct
acacctgaaaatattcctgaaactatcgaagtagatatcagtgaattacaagttaatgac
agcttagcagtttctgatattaaaatctctggtgatttcacaatcgaaaataatccagaa
gattctatcgtaacagtagttcctccaacagatgaaccttctgaagaagaagttgaagct
atggaaggcgaatcagcaactgaagaaccagaagttgttggtgaagacaaagaagacgat
gaagaagaaataaagaagactaa
```

Sequence 2502
MVSDYQFDPLKNQITHIDFLAINMSEERTVEVQVQLVGEAVGAKEGGVVEQPLFNLEVTA
TPENIPETIEVDISELQVNDSLAVSDIKISGDFTIENNPEDSIVTVVPPTDEPSEEEVEA
MEGESATEEPEVVGEDKEDDEEENKED*

Sequence 2503
Contig_0766_pos_5387_0,
is similar to (with p-value 4.0e-49)
>sp:sp|P37470|SP5C_BACSU PROBABLE PEPTIDYL-TRNA HYDROLASE (E
C 3.1.1.29) (PTH) (STAGE V SPORULATION PROTEIN C). >gp:gp|D2
6185|BAC180K_116 B. subtilis DNA, 180 kilobase region of rep
lication origin. NID: g467326. >gp:gp|Z99104|BSUB0001_53 Bac
illus subtilis complete genome (section 1 of 21): from 1 to
213080. NID: g2632267.
```
gtggaggtaacaataatgaaatgcattgtcggtcttggcaacattggtaaacgttttgaa
ttaacaagacataatattggtttcgaagttgtcgatgatattctagaacgccaccaattt
acttagacaaacaaaaatttaaaggtgcatatactattgaacgtttaaacggcgaaaaa
gtatcattattgagccaatgaccatgatgaacttatctggtcaagctgtagcccttta
atggattattataatgtcgatgttgaagatttgatcgtttatatgacgatttagattta
gaacaaggacaagtgcgtctgcgccaaaaggggagtgcaggcggtcataatggtatgaaa
tcgataattaaaatgcttggtacagatcaatttaaacgtattcgaattggtgttggccgt
ccaacaaatgggatgtctgttccggactatgtttacaaaaattttcaaaagaagaaatg
atcattatggaaaaggtaattgaacattctgcaagagctgtagaatctttttattgaaagt
tctcgttttgatcatgttatgaatgaa
```

Sequence 2504
VEVTIMKCIVGLGNIGKRFELTRHNIGFEVVDDILERHQFTLDKQKFKGAYTIERLNGEK
VLFIEPMTMMNLSGQAVAPLMDYYNVDVEDLIVLYDDLDLEQGQVRLRQKGSAGGHNGMK
SIIKMLGTDQFKRIRIGVGRPTNGMSVPDYVLQKFSKEEMIIMEKVIEHSARAVESFIES
SRFDHVMNE

Sequence 2505
Contig_0766_pos_1682_1317,
putative peptide of unknown function
```
atgtatagtgaaaaagaaatcatacgaaaagtcgaaagtttagcagagaaaattggaaaa
ctagaagttgttcaagattatcataatgtagaaaaacaaattcataataatcaagcaata
aaacaaaagatgaatcgtttgaaagcgcaacaaaaacaatcggttaattttcaaaattat
ggaaaacaaaatgcactcgagcaatctgaagttaaaattcagaatctaaaagatgaaatt
aatgaattacctattgttgaagaatttcgttcagcacaatatgaagcgaatgatttacta
caaatgatggtcaaaacaatggaagatagactcaatgaatataataaaaaagaacacaat
gaataa
```

Sequence 2506
MYSEKEIIRKVESLAEKIGKLEVVQDYHNVEKQIHNNQAIKQKMNRLKAQQKQSVNFQNY
GKQNALEQSEVKIQNLKDEINELPIVEEFRSAQYEANDLLQMMVKTMEDRLNEYNKKEHN
E*

Sequence 2507
Contig_0766_pos_1293_799,
putative peptide of unknown function

```
atgaatattagaaagttgaccatcacagcattcttaatagctattaatgtcgtgttaagt
agtttaattgtcattcctttaggtccaattaaagccgcacccgttcaacattttgtaaat
gtattatgtgctgtatttgttggaccatggtatggtttagcgcaagcttttatttcttct
gtactacgaatttcatttggaactggaagcgcatttgcttttccaggaagtatgataggt
gtcttactttccagtctgttttatatgtataggaagcatatttttatggcttcagttggt
gaagtattagggactggtgttattggtagtttaatgtgtatacctttagcatggttttta
ggacttcaagatttctttattaaaccattaatgcttatgttcatagtatcaagtttttatt
ggggctttaattagttatatattgcttattattttgaaaagaagaggtttactagataga
tttaacaaaaattaa
```

Sequence 2508

```
MNIRKLTITAFLIAINVVLSSLIVIPLGPIKAAPVQHFVNVLCAVFVGPWYGLAQAFISS
VLRISFGTGSAFAFPGSMIGVLLSSLFYMYRKHIFMASVGEVLGTGVIGSLMCIPLAWFL
GLQDFFIKPLMLMFIVSSFIGALISYILLIILKRRGLLDRFNKN*
```

Sequence 2509
Contig_0766_pos_0_613,
is similar to (with p-value 4.0e-50)
>sp:sp|P49849|MUTS_BACSU DNA MISMATCH REPAIR PROTEIN MUTS. >
gp:gp|U27343|BSU27343_2 Bacillus subtilis spore coat protein
 (cotE) gene, partial cds, and mismatch repair recognition p
 roteins (mutS) and (mutL) genes, complete cds. NID: gl002518

```
gtgaaaatagaaatggctaacattacaccaatgatgcaacaatatttaaagataaaatct
gaatatgatgattgtttgctattttttagactcggagatttctatgaaatgttctttgat
gatgctaaagaagcatcaagagtacttgaaataacattgacgaaaagagatgctaaaaaa
gaaaatcctattccgatgtgtggcgtaccatatcattctgctgataattacattgaaaca
ttgattaataatgggtataaggtcgctatatgtgaacaaatggaagatccaaagcaaaca
aaaggaatggttagaagagaagttgtaagaatcatcacaccagaactgttatggatcaa
aatggtatggatgaaaagaaaaataattatatttttaagttttatcgaaaatgaagaattt
ggattatgctattgtgatgtttctacaggcgaactcaaagtaactcatttcaaagataca
gcaaccttgcttaatgagattacaacaattaatcccaatgaaatcgtcataaagcaagct
ctatctgaagaattaaaaagacaaatcaacatgataactgagacgattactgttcgcgaa
gatataTATGTTT
```

Sequence 2510

```
VKIEMANITPMMQQYLKIKSEYDDCLLFFRLGDFYEMFFDDAKEASRVLEITLTKRDAKK
ENPITMCCVPYHSADNYIETLINNGYKVAICEQMEDPKQTKGMVRREVVRIITPGTVMDQ
NGMDEKKNNYILSFIENEEFGLCYCDVSTGELKVTHFKDTATLLNEITTINPNEIVIKQA
LSEELKRQINMITETITVREDIYVX
```

Sequence 2511
Contig_0767_pos_5389_5703,
is similar to (with p-value 1.0e-24)
>gp:gp|AF006075|AF006075_3 Bacillus subtilis acetoin dehydro
genase enzyme system gene cluster, ribosomal protein L6-like
 protein gene, partial cds, TPP-dependent acetoin dehydrogen
ase, E1 alpha-subunit (acoA), TPP-dependent acetoin dehydrog
enase, E1 beta-subunit (acoB), dihydrolipoamide acetyltransf
erase (acoC) and dihydrolipoamide dehydrogenase (acoL) genes
, complete cds, and regulatory protein (acoR) gene, partial
cds. NID: g2957145. >gp:gp|Z99108|BSUB0005_76 Bacillus subti
lis complete genome (section 5 of 21): from 802821 to 101125
0. NID: g2633055. >gp:gp|D78509|D78509_4 Bacillus subtilis Y
fjG-YfjR genes, complete cds. NID: g2780390.
atgagttcagcaattggacgtagccctgtcgcagcagctcctactgccgtgctcaatgta
atgtgttcagcaattggtgtatcgattacacgtttacgactatattttttgcaagtcct
tttgttataaccaaatacaccaccgaatgtatcgtcatctttgatgtggtctacttatgca
ccacctgagacatcagtaccaattaaaatgacatcctcatctttttccatagattggtca
```

atagcttcattaatagcccccataaatgttaacttgcgttcttcactcatcgtaaaaacc
tccttatgtggttag

Sequence 2512
5   MSSAIGRSPVAAAPTAVLNVMCSAIGVSITRLRLYFFASPFVTPNTPPNVSSSLMWSTFA
    PPETSVPIKMTSSSFSIDWSIASLIAPINVNLRSSLIVKTSLCG*

Sequence 2513
Contig_0767_pos_5685_4645,
10  is similar to (with p-value 0.0e+00)
    >gp:gp|AF006075|AF006075_3 Bacillus subtilis acetoin dehydro
    genase enzyme system gene cluster, ribosomal protein L6-like
     protein gene, partial cds, TPP-dependent acetoin dehydrogen
    ase, E1 alpha-subunit (acoA), TPP-dependent acetoin dehydrog
15  enase, E1 beta-subunit (acoB), dihydrolipoamide acetyltransf
    erase (acoC) and dihydrolipoamide dehydrogenase (acoL) genes
    , complete cds, and regulatory protein (acoR) gene, partial
    cds. NID: g2957145. >gp:gp|Z99108|BSUB0005_76 Bacillus subti
    lis complete genome (section 5 of 21): from 802821 to 101125
20  0. NID: g2633055. >gp:gp|D78509|D78509_4 Bacillus subtilis Y
    fjG-YfjR genes, complete cds. NID: g2780390.
    atgagtgaagaacgcaagttaacatttatgggggctattaatgaagctattgaccaatct
    atggaaaaagatgaggatgtcattttaattggtactgatgtctcaggtggtgcaaaagta
    gaccacatcaaagatgacgatacattcggtggtgtatttggtgtaacaaaaggacttgca
25  aaaaaatatagtcgtaaacgtgtaatcgatacaccaattgctgaacacattacattgagc
    acggcagtaggagctgctgcgacagggctacgtccaattgctgaactcatgttcaacgac
    tttattggatttggtttagatccaattttaaatcaaggggcaaaaatgagatatatgttt
    ggtggaaaagccaaaatcccactagttgtacgtactcttcatggagcaggggcaagcgct
    gctgcacagcactctcagtctttatataatatgtttgcagcaattccaggagttaaagtt
30  gttgttccatctaatccatatgatgcgaagggtctactgatgtcagctattcaagaggac
    aatcttgttgtctttcagaagatataaaacattattaggacaaaaaggtaatgttcctgaa
    gaaccttatactatagaaattggtaaagccaatgtgacgcgtgaaggtgacgatttaaca
    attgtggctattggaaaaatggtagctgtagcggaagaaactgctgaaaaacttgcagaa
    gaccaagtatcagttgaggtcatcgatttacgctcagtgtcaccatgggatcaagaaaca
35  gttttagattctgtgaagaaaacgggtcgcttaattgttattgacgaatctaatccacag
    tgtaacattgctggagacgttgcttcagtgattggagatgtaggatttgattacttagat
    ggtccaattaagaaagtgaccgcaccagacactcctgtaccatttgcagcgaacttagag
    gcggcatatatgccgaatgctgataaggtattagacattgcatctgaattaattgatgat
    ttaaaaaaggctaacgcatag
40
    Sequence 2514
    MSEERKLTFMGAINEAIDQSMEKDEDVILIGTDVSGGAKVDHIKDDDTFGGVFGVTKGLA
    KKYSRKRVIDTPIAEHITLSTAVGAAATGLRPIAELMFNDFIGFGLDPILNQGAKMRYMF
    GGKAKIPLVVRTLHGAGASAAAQHSQSLYNMFAAIPGVKVVVPSNPYDAKGLLMSAIQED
45  NLVVFSEDKTLLGQKGNVPEEPYTIEIGKANVTREGDDLTIVAIGKMVAVAEETAEKLAE
    DQVSVEVIDLRSVSPWDQETVLDSVKKTGRLIVIDESNPQCNIAGDVASVIGDVGFDYLD
    GPIKKVTAPDTPVPFAANLEAAYMPNADKVLDIASELIDDLKKANA*

    Sequence 2515
50  Contig_0767_pos_4613_3354,
    is similar to (with p-value 3.0e-72)
    >gp:gp|Z99108|BSUB0005_77 Bacillus subtilis complete genome
    (section 5 of 21): from 802821 to 1011250. NID: g2633055. >g
    p:gp|D78509|D78509_3 Bacillus subtilis YfjG-YfjR genes, comp
55  lete cds. NID: g2780390.
    atgccaaagcttggaatgacaatgaaagagggaactgttgaagagtggtttaaatcagag
    ggtgacaccgtaaaacaaggagagagtattgttacaataagctctgaaaaattaaccaac
    gatgttgaagcgccggcgagtgggacattgttagaaattaaagtgcaagccggagaagat
    gcagaggttaaagcggtattaggtataattggagaagaagggggaagctattgataaagat

gaagatgatttagcatcagaaaaagtaaaagaagacaacgagcatgagaaggaaacgcaa
gaagttaaagatacatcacaacagtcttccgataataaagataattcgcctaaaagcgca
gcacgagaaagaatctttatctcacccctcgcacgtaatatggctgaggataaaggatta
gacattaacaagataaaaggcacaggcggtaatcatcgtattacaaaactagatattcaa
cgtgttgaagcaaatgggtacgactatgctagtgatacgacatctaatgaagatacaagt
catgttccaacacagactgtggatacaagtgcgattggtgaaggattgaatcctatgcgt
caacgtattgctcaaaacatgagacaaagtcttaatagtactgctcaattaacattacat
cgtaaggttgatgcggatcgcttgctagatttcaaagacagattagctacggaacttaaa
caagcagatcaagatgttaaattaactgttactacattattagctaaagcagtagtgctt
gcacttaaagaatatggggcaatgaatgctcgctatgaacaaggcgagttaactgagtat
gaagatgttcatttaggaatcgcaacgtctctagatgaaggccttatggtgccagtgatt
aatcatgcagatacaaaaagtatcggcactttagcccatgaaattaaatcatcggctgag
gctgttcgggaaggaaacacaggagcagtacaattagagggagcaacatttacaattact
aatatgggtgctagtggtatagaatactttacaccaattttaaatttaggtgaaacaggt
attctaggcgttggtgctttaactaaagaagtcgtgctagaagcggataacattaaacaa
gtttcaaaaattcctttaagcttgacatttgatcatcaaattttagatggtgcaggtgcg
gccgattttcttaaagtactagctaaatatatcgaaaacccttatttattaatgttatag

Sequence 2516
MPKLGMTMKEGTVEEWFKSEGDTVKQGESIVTISSEKLTNDVEAPASGTLLEIKVQAGED
AEVKAVLGIIGEEGEAIDKDEDDLASEKVKEDNEHEKETQEVKDTSQQSSDNKDNSPKSA
ARERIFISPLARNMAEDKGLDINKIKGTGGNHRITKLDIQRVEANGYDYASDTTSNEDTS
HVPTQTVDTSAIGEGLNPMRQRIAQNMRQSLNSTAQLTLHRKVDADRLLDFKDRLATELK
QADQDVKLTVTTLLAKAVVLALKEYGAMNARYEQGELTEYEDVHLGIATSLDEGLMVPVI
NHADTKSIGTLAHEIKSSAEAVREGNTGAVQLEGATFTITNMGASGIEYFTPILNLGETG
ILGVGALTKEVVLEADNIKQVSKIPLSLTFDHQILDGAGAADFLKVLAKYIENPYLLML*

Sequence 2517
Contig_0767_pos_2671_2159,
putative peptide of unknown function
atgaaaaggtttgcaaaagcatttgtcgtaagtggtattactttaggtgcagttttaggt
ttaaacgtaacagagcataatggtgtatctaatgaagcaaaggcacaaacagcacacagt
tactggtataaatataatggttatactgcatcgggtggcgactttgtacttagcaattca
ttttatcaaggtttaaaagctggaaacgttacatttaatggtattaaggtaaatcaaaaa
tatgaatctaagactgctactaaaaaaatatacgatcagacatttcaacaaattaatgga
aataaagcaaacaacgtacaatttaaaattgcttccagaactgttactttagatcaagtt
aaacaaaagtatggaaaaaattataattatcagccgtcattatctaaaaacaaaacaagt
aacacagatggcttgtacggttatcaagtcggaaaaggaaacatcgttttccacgttaaa
gatgtgtatgtcacaagtgctacattgtcataa

Sequence 2518
MKRFAKAFVVSGITLGAVLGLNVTEHNGVSNEAKAQTAHSYWYKYNGYTASGGDFVLSNS
FYQGLKAGNVTFNGIKVNQKYESKTATKKIYDQTFQQINGNKANNVQFKIASRTVTLDQV
KQKYGKNYNYQPSLSKNKTSKTDGLYGYQVGKGNIVFHVKDGYVTSATLS*

Sequence 2519
Contig_0767_pos_1876_1568,
putative peptide of unknown function
atgagtgagaatatcattaataataaaatgatagagaatggtagacctgtaataactgat
gcagtttgtaaacttttaagaccacctgcaattgtcatagcaactgaaatcgccccaata
agtagaccccataggactttatgttttaaagttggattgatagaaccacctgtagccata
cctgcaacaatatgcgtagttgaatcagcacttgttacaataaagatgaaaataagtaca
atagctaaggaactagtgacttccgacaatggaaaatgtgataataattcaaacaaggca
acagtataa

Sequence 2520
MSENIINNKMIENGRPVITDAVCKLLRPPAIVIATEIAPISRPHRTLCFKVGLIEPPVAI

PATICVVESALVTIKMKISTIAKELVTSDNGKCDNNSNKATV*

Sequence 2521
Contig_0768_pos_2298_2642,
putative peptide of unknown function
gtgacaaaccggaggaaggtgggggatgacgtcaaatcatcatgcccccttatgatttgggc
tacacacgtgctacaatggacaatacaaagggtagcgaaaccgcgaggtcaagcaaatcc
cataaagttgttctcagttcggattgtagtctgcaactcgactatatgaagctggaatcg
ctagtaatcgtagatcagcatgctacggtgaatacgttcccgggtcttgtacacaccgcc
cgtcacaccacgagagtttgtaacacccgaagccggtggagtaaccatttggagctagcc
gtcgaaggtgggacaaatgattggggtgaagtcgtaacaaggtag

Sequence 2522
VTNRRKVGMTSNHHAPYDLGYTRATMDNTKGSETARSSKSHKVVLSSDCSLQLDYMKLES
LVIVDQHATVNTFPGLVHTARHTTRVCNTRSRWSNHLELAVEGGTNDWGEVVTR*

Sequence 2523
Contig_0768_pos_7279_6914,
is similar to (with p-value 1.0e-27)
>gp:gp|U96108|SCU96108_3 Staphylococcus carnosus (3R)-hydrox
ymyristoyl acyl carrier protein dehydrase homolog (fabZ) gen
e, partial cds, YwpF homolog, single-strand binding protein
homolog (ssb), SceD precursor (sceD), SceA precursor (sceA)
and SceE precursor (sceE) genes, complete cds, and TenA homo
log (tenA) gene, partial cds. NID: g2735509.
atgacgaaagatgctcaaatctatgaaaaggaggataacaaaatagcaacattttgcgtt
gctacagaaagaaactacaaagatgatattaatgaaatttctacagattatttactttgt
aaagctttcggtaaaactgccactaatattgaaaagtacactagtcaaggtactttagta
ggtataactggacaaatgcgttcaagaaaatatgaaaaagaaggccaaacacattttgtt
acagaattatacgttgaaacaataaaattcatgtccccaaaaaataaaaacaatgaaact
ccctctgataatcaatttgaaaacaacacttatcaacctgatgatttagaaataattcat
att.taa

Sequence 2524
MTKDAQIYEKEDNKIATFCVATERNYKDDINEISTDYLLCKAFGKTATNIEKYTSQGTLV
GITGQMRSRKYEKEGQTHFVTELYVETIKFMSPKNKNNETPSDNQFENNTYQPDDLEIIH
I*

Sequence 2525
Contig_0768_pos_6526_5867,
is similar to (with p-value 5.0e-49)
>gp:gp|U96108|SCU96108_4 Staphylococcus carnosus (3R)-hydrox
ymyristoyl acyl carrier protein dehydrase homolog (fabZ) gen
e, partial cds, YwpF homolog, single-strand binding protein
homolog (ssb), SceD precursor (sceD), SceA precursor (sceA)
and SceE precursor (sceE) genes, complete cds, and TenA homo
log (tenA) gene, partial cds. NID: g2735509.
atgaaaaaaacattagttgcatcatcattagctataggactaggcgttgtagcaggtaac
gcaggtcatgacgcacatgcaagtgagactacaaatgttgataaagcagagttagctcaa
aaagcgttaactaatgatcaatcactaaatgaaagccctgttcaagaaggtgcttataat
attaattttgattacaatggtaattcatatcattttgaatctgatggttctacttgjagt
tggagct.tgaatcaacaaacaatgctacgcaacctgttcaaccaagccaatctcagta
gctacacaacaacaacctgtacaagtatcagcaccacaaaatgagcaaacagcacaacca
caaacaaatcaacatctacaagtcaaacttcttcaagtaaagcttcaagtggttcatca
gtaaatgtaaattcacatttacaacaaattgctcaacgtgaatcaggtggcgatattcat
gcaataaacccatcttcaggtgcagcaggtaaatatcaattcttacaatcaacttgggat
tctgttgctccaagtcaatataaaggtgtttcacctgcaaaagctccagaaagcgtacaa
gaccgagcagcagtaaaattatataatactggtggcccaggtcattgggtaactgcataa

Sequence 2526
MKKTLVASSLAIGLGVVAGNAGHDAHASETTNVDKAELAQKALTNDQSLNESPVQEGAYN
INFDYNGNSYHFESDGSTWSWSYESTNNATQPVQPSQSQVATQQQPVQVSAPQNECTAQP
QTKSTSTSQTSSSKASSGSSVNVNSHLQQIAQRESGGDIHAINPSSGAAGKYQFLQSTWD
SVAPSQYKGVSPAKAPESVQDRAAVKLYNTGGPGHWVTA*

Sequence 2527
Contig_0768_pos_5390_4893,
is similar to (with p-value 3.0e-44)
>gp:gp|U96108|SCU96108_7 Staphylococcus carnosus (3R)-hydrox
ymyristoyl acyl carrier protein dehydrase homolog (fabZ) gen
e, partial cds, YwpF homolog, single-strand binding protein
homolog (ssb), SceD precursor (sceD), SceA precursor (sceA)
and SceE precursor (sceE) genes, complete cds, and TenA homo
log (tenA) gene, partial cds. NID: g2735509.
atggaagatgttaaatttctagtagaacaaattgagtttatgttagaaggagaagttgaa
gcgcatgaggttctagctgatttcattaatgaaccatatgaagaaatagtaaaagaaaaa
gtatggccaccaagtggtgatcattatattaaacatatgtacttcaatgcatttgcacgt
gaaaatgcagccttcacgattgcagcgatggcaccctgtccatacgtctacgctgtcatt
ggtaaacgtgcgatggaagatcccaaattaaataaagaatcagtgacttctaaatggttt
caattttatagtactgaaatggacgaacttgttgatgtgttcgatcaattgatggaccgt
ttaactaaacattgtagtgagacagaaaaaaaagagattaaagaaaatttcttgcaaagt
actattcatgagagacatttcttcaatatggcatatattaatgaaaaatgggaatatggg
ggaaataacaatgaataa

Sequence 2528
MEDVKFLVEQIEFMLEGEVEAHEVLADFINEPYEEIVKEKVWPPSGDHYIKHMYFNAFAR
ENAAFTIAAMAPCPYVYAVIGKRAMEDPKLNKESVTSKWFQFYSTEMDELVDVFDQLMDR
LTKHCSETEKKEIKENFLQSTIHERHFFNMAYINEKWEYGGNNNE*

Sequence 2529
Contig_0768_pos_4639_4079,
is similar to (with p-value 8.0e-41)
>sp:sp|P44697|THID_HAEIN PHOSPHOMETHYLPYRIMIDINE KINASE (EC
2.7.4.7) (HMP-PHOSPHATE KINASE) (HMP-P KINASE). >pir:pir|I64
151|I64151 hypothetical protein HI0416 - Haemophilus influen
zae (strain Rd KW20) >gp:gp|U32725|U32725_3 Haemophilus infl
uenzae Rd section 40 of 163 of the complete genome. NID: g15
73387.
atggagttgattcaaagttacttaaaggaacattccaatattccttatgtgattgatcca
gttatgcttgccaaaagtggtgattcattgatggatgaaaatactaaaaatcatttgcaa
tcaacattattacctttggctgatgttgtcactccgaatatacctgaagctgaggaaatt
actggtattaaaattaatgatgaagaaagcatacgtaaagcaggtcaaatctttattaat
gaaattggtagtaagggtgttgtaattaagggagggcattcagccgatttaaataatgct
aaagattttcttttttactaagaatgaaacgtatacctttgagaacaaacgctttgatact
aaacatactcatggaactggttgtacttttcagcagttattacagcagaattagctaaa
ggtcgttccataaaagatgcagttaaaaaagcaaaagagtttatttcattaagtattgaa
cataccccagaaattggcaagggggagaggacctgtaaatcattttgcttatatgaagaaa
gtaggtttagatgatgaataa

Sequence 2530
MELIQSYLKEHSNIPYVIDPVMLAKSGDSLMDENTKNHLQSTLLPLADVVTPNIPEAEEI
TGIKINDEESIRKAGQIFINEIGSKGVVIKGGHSADLNNAKDFLFTKNETYTFENKRFDT
KHTHGTGCTFSAVITAELAKGRSIKDAVKKAKEFISLSIEHTPEIGKGRGPVNHFAYMKK
VGLDDE*

Sequence 2531
Contig_0769_pos_1300_1833,

is similar to (with p-value 1.0e-69)
>sp:sp|P37948|GLPT_BACSU GLYCEROL-3-PHOSPHATE TRANSPORTER (G
-3-P TRANSPORTER) (G-3-P PERMEASE). >pir:pir|S37250|S37250 g
lycerol-3-phosphate transport protein - Bacillus subtilis >g
5    p:gp|Z26522|BSGLPTQ_1 B.subtilis glpT and glpQ genes for gly
cerol 3-phosphate permease and glycerophosphoryl diester pho
sphodiesterase. NID: g403371. >gp:gp|Z99105|BSUB0002_43 Baci
llus subtilis complete genome (section 2 of 21): from 194651
    to 415810. NID: g2632457. >gp:gp|AB006424|AB006424_44 Bacil
10   lus subtilis genomic DNA, 70 kb region between 17 and 23 deg
ree. NID: g3599592.
atgtttaatttcttaaaaccagctagacatatcaaatcattgccatcagaaaaagtagat
gatacgtataaaagactacgctttcaagtctttttaggaatatttatcggttatgctggt
tactatttattaagaaagaacttctctttagcaatgccttcattaattgagcaaggcttt
15   agcatacgtggaattaggtattgcattatctgcagtatctatcgcatatggctttagcaaa
tttgtaatgggcactgtcagcgatcgaagtaatgctcggatgttcttaactttaggttta
gtattgacagcaattattaacttattattaggatttattccattctttacttcaagcata
actattatgtttatcatgctgtttttagttggatggttccaaggaatgggctggccacca
tctggacgtgtgttagttcattggtttagtgtcagcgaacgtggaagcaaaacgtcaata
20   tggaatgtagcacataatgtaggcggaggtttaatggcacctattgctacgtag


Sequence 2532
MFNFLKPARHIKSLPSEKVDDTYKRLRFQVFLGIFIGYAGYYLLRKNFSLAMPSLIEQGF
SKGELGIALSAVSIAYGFSKFVMGTVSDRSNARMFLTLGLVLTAIINLLLGFIPFFTSSI
25   TIMFIMLFLVGWFQGMGWPPSGRVLVHWFSVSERGSKTSIWNVAHNVGGGLMAPIAT*


Sequence 2533
Contig_0769_pos_1843_2658,
is similar to (with p-value 2.0e-72)
30   >sp:sp|P37948|GLPT_BACSU GLYCEROL-3-PHOSPHATE TRANSPORTER (G
-3-P TRANSPORTER) (G-3-P PERMEASE). >pir:pir|S37250|S37250 g
lycerol-3-phosphate transport protein - Bacillus subtilis >g
p:gp|Z26522|BSGLPTQ_1 B.subtilis glpT and glpQ genes for gly
cerol 3-phosphate permease and glycerophosphoryl diester pho
35   sphodiesterase. NID: g403371. >gp:gp|Z99105|BSUB0002_43 Baci
llus subtilis complete genome (section 2 of 21): from 194651
    to 415810. NID: g2632457. >gp:gp|AB006424|AB006424_44 Bacil
lus subtilis genomic DNA, 70 kb region between 17 and 23 deg
ree. NID: g3599592.
40   atgactgcattatacaacttcggttatttaaaagggtttgaaggtgtctttatataccct
gcactattggttatcattattgccatcttgtcttacatactaattagagatacaccacaa
tctcagggtttaccaccaattgagcagtataaaaatgattatgccacttcaactaaacaa
acaattgaaacagaactaactacgaaagaaatattatttaaatatgtacttaataacaaa
tgggtatgggcgattgcttttacaaacattttcgtttattttgtgcgttatggtgttttg
45   gactgggctccgacatacttaagtgaggaaaagcattttgatttaagtgcttcaggttgg
gcttacttcttatacgaatgggcaggaattccgggcacgctactctgtggttatctatct
gacaaactatttaaaggtcgccgtggtccagcaggcttcttctttatgttaggcgtaaca
atctttatccttatatattggttaaatccaccaggtcacgcatggttagataatctttca
ttaattggtattggtttcttaatttatggtccagtcatgttaattggtttacaagcgtta
50   gattatgttcctaaaaaagcagctggcactgcggctggtctaactggtctatttggatac
ttattcggtgcagttatggctaatattgtattagggtttgtagttcaacattttggatgg
catattggctttgtgttattaacagtcatcagcatactcgctatgttatgtttcatttta
acttggaataaacgtggtcaagaacaaatcgactag


55   Sequence 2534
MTALYNFGYLKGFEGVFIYPALLVIIIAILSYILIRDTPQSQGLPPIEQYKNDYATSTKQ
TIETELTTKEILFKYVLNNKWVWAIAFTNIFVYFVRYGVLDWAPTYLSEEKHFDLSASGW
AYFLYEWAGIPGTLLCGYLSDKLFKGRRGPAGFFFMLGVTIFILIYWLNPPGHAWLDNLS
LIGIGFLIYGPVMLIGLQALDYVPKKAAGTAAGLTGLFGYLFGAVMANIVLGFVVQHFGW

HIGFVLLTVISILAMLCFILTWNKRGQEQID*

Sequence 2535
Contig_0769_pos_4799_5716,
putative peptide of unknown function
atggttatttcaaagaaattagcatttatcactataggtatcttaatatgtattctttta
agttcacctctgatcgatcaacagatttcaaattatttcatgaatcaagattcaatttttc
ggtactttatttcaaaactacggtttatttccacctacgctcatattaattatttcaact
gtaattttaaattattatattttaactacatttcaaaataaattagcaaaaacgttaact
ttgattttatcttttatctttacgttgataaaaaccaatgaattcgtgtcagaaaccgct
caatatatgttatctacatcagaaaacatcaaaaatcataagccaatgggtatggcaaat
aatgaaggaaacgctggaaatgctttatctttaggaatgagttttttttatttctttgatt
atcattataattatcactattatttgttatcagttttggttaaaacatacaaataaccaa
gaacttgatcatttattcaaagtttcacttattagctttatgattttatttattggatta
gaactcgttgatagcttgaaacatttatggggacgttttagaccatatgaaatcactgac
aaagctggacatttcactcactggcttacaataaatggaaacactggacatagttccttc
ccttcaggacacactggaaatggtgcatttctaatgtttatcgcattttactttaaaaaa
ttacgcactcaaaaaatagtgtttagtatcggattatgttatagtattttaatggcattg
agcagaattcgtattggtgcacatttcactagtgatgtaacaatgagcttgcttattatg
ttttcactcatggttattgcagattttatcattaatcaagtcatatctatacataaaaac
aagctaagtaaagtctag

Sequence 2536
MVISKKLAFITIGILICILLSSPLIDQQISNYFMNQDSIFGTLFQNYGLFPPTLILIIST
VILNYYILTTFQNKLAKTLTLILSFIFTLIKTNEFVSETAQYMLSTSENIKNHKPMGMAN
NEGNAGNALSLGMSFFISLIIIIIITIICYQFWLKHTNNQELDHLFKVSLISFMILFIGL
ELVDSLFHLWGRFRPYEITDKAGHFTHWLTINGNTGHSSFPSGHTGNGAFLMFIAFYFKK
LRTQKIVFSIGLCYSILMALSRIRIGAHFTSDVTMSLLIMFSLMVIADFIINQVISIHKN
KLSKV*

Sequence 2537
Contig_0769_pos_7866_7003,
putative peptide of unknown function
atgccgttaatcttaggtatcgaaaaaggttggtttaaagaagaatcattagaaattgag
atgattgaaccaaaggaacactttgacgcactagacgagagtgaaaatggttcaatggat
atcgcgattactgaacctattcatctggttcaagatagagctaaagaacaaaaagtcatc
gggtttgcgagatatcttcacactaatggaggtattatgtataacaaagataaaaaatatc
gctcgccccaaagatttaatcggtaaacggcttcaatatcccggtgctcccggtccaggt
ggtattgctatggctaaaacgatgattgaagctgatggtggtacatttgaagaaggtgac
atcacaccagttaatcatagttttttatcatactgatgcacttttaactgataaagctgat
gctgctacactcattttcgaaaattttgaaattcttgaagctagaaatcaaggacttaat
gtagattattttccacttaaaaattataatgtacccgatttttgtcaactcattttcatt
acaacacctgaagtattaaatacgagtgaggttaaactcaaaaagttcctaaaaatcatt
caaaaaacaattcattacatcaatcatcatttagaaagtgcaattgaaatttactcaacg
tatactcaaaccgatatctctaaccaattaaacaaagatacaattgaagcaacagctaaa
tgttttacaaatgatttgtctatgagttccgactactacaatgatttacagttgtggctt
aaagaagtcaatgatattaaagacacgattaacccaactttatattttacaaatcaacta
ctattcagtcgatatacaaaataa

Sequence 2538
MPLILGIEKGWFKEESLEIEMIEPKEHFDALDEIENGSMDIAITEPIHLVQDRAKEQKVI
GFARYLHTNGGIMYNKDKNIARPKDLIGKRLQYPGAPGPGGIAMAKTMIEADGGTFEEGD
ITPVNHSFYHTDALLTDKADAATLIFENFEILEARNQGLNVDYFPLKNYNVPDFCQLIFI
TTPEVLNTSEVKLKKFLKIIQKTIHYINHHLESAIEIYSTYTQTDISNQLNKDTIEATAK
CFTNDLSMSSDYYNDLQLWLKEVNDIKDTINPTLYFTNQLLFSRYTK*

Sequence 2539
Contig_0769_pos_6570_6241,
putative peptide of unknown function

```
atgggtttgaagagaaattggaggttgagaaagatgtatgtagagagaaagcctgcactt
tatatcgaagatttacgagacgaattcaaaaatagcttaaaacattttaaagatgatgat
gaagcatttaatacattagttggatttgtggaactggaccacctatattcttcagcactt
aaagaaataagtacaaaattaagtattttagatgataactttaactatcaatataaacac
```
5
```
aatccaattcatcatatggaacgacgagtaaaagagatgcacagtttagttaaaaagttg
aatcgtaaaggatttgaagtaagtgcataa
```

Sequence 2540
MGLKRNWRLRKMYVERKPALYIEDLRDEFKNSLKHFKDDDEAFNTLVGFVELDHLYSSAL
10  KEISTKLSILDDNFNYQYKHNPIHHMERRVKEMHSLVKKLNRKGFEVSA*

Sequence 2541
Contig_0769_pos_4499_3618,
putative peptide of unknown function
15
```
atgtatatgattaagcaaaattttacaactaaagtttatccacttaaacaatttgaggga
gaatatagtgaagttgcctacgcaggagatagccaagctgatgatgtgattgtatttatt
catggtgcacttttaacatataaaattatgacgatgttcgaaccttacttcagagactac
aaattaatatttattaattgtccaagtcgtggtagaagttcagatttagatcgtgacaca
catacattagatgattacgctgcacgtatatatgatgtattaacgcaaattgttaaggag
```
20
```
caacaaataaaagaactgagtattgtcggttattcaatgggaggaatgattgcgacacgg
ttacttaagtataatacattaccagtctctcatcttatttatttacatagcgcagcgaaa
attactccagatgcaagtatgttagcacgattattcactagtgagagtaagagagcagtg
ttaaaagatgaaattaaggcagtgaaaaatcttcctcaatatactagtaaaactgatt
tatgcacaaaaggaaaacgcacttgatttggtacaatttattgcacctattaaaactata
```
25
```
attacggatatgatttacacgattaatacagattatttaccagatatcgatgagattaaa
caatttccgaaaatattatttatgtctggaaaagaagatcaaattattccttatacggat
tctcaagctacgttagaaaagtttaaggcgttaggtggagaaactaaagaagttatttat
ccaggaattggtcatatcgatttcccaagtgttttagaaacccaatcggatggacaaact
ggcgtggtagatgaaattaaagcgtggatttcaaaaaaataa
```
30

Sequence 2542
MYMIKQNFTTKVYPLKQFEGEYSEVAYAGDSQADDVIVFIHGALLTYKIMTMFEPYFRDY
KLIFINCPSRGRSSDLDRDTHTLDDYAARIYDVLTQIVKEQQIKELSIVGYSMGGMIATR
LLKYNTLPVSHLIYLHSAAKITPDASMLARLFTSESKRAVLKDEIKAVKNLPQYILDKTI
35  YAQKENALDLVQFIAPIKTIITDMIYTINTDYLPDIDEIKQFPKILFMSGKEDQIIPYTD
SQATLEKFKALGGETKEVIYPGIGHIDFPSVLETQSDGQTGVVDEIKAWISKK*

Sequence 2543
Contig_0769_pos_3394_2816,
40  putative peptide of unknown function
```
gtggagaaacttgaactgattttactacttaaggaaataggttgcacaattaaagaaatt
aaagtgttattaaaagatgattcatcaatgaaatctctttacacgtttcttcaaatcaaa
aagcatgaaatacaacagtcaataacagataaagaaacaaggtatgtaaaatagagcaa
attcaacgctatgttcatcaaaattcaatttctccaattcattatttacaagatatagct
```
45
```
atttatatggaagaatctcatagactcaaaggtgtcagaaagaaattatggttaagtata
gctctgattggttcattacaatatggtggtttgatcacctcaattgtaactcaaaggaaa
aagcctttcttgtcgatgatgcctgtggtagctatgtattcactttggttaacgaaaaag
tataaaaagaatgtttcatatgtctgtcctaattgccaccatgtgtttaatcctagtgtc
attcattttgtaacagcatcacatacacctaaacaagaaagcttcaatgtcctgattgt
```
50  cacgaaatgcattactgtgtagaaattgctaagctatga

Sequence 2544
VEKLELILLLKEIGCTIKEIKVLLKDDSSMKSLYTFLQIKKHEIQQSITDKENKVCKIEQ
IQRYVHQNSISPIHYLQDIAIYMEESHRLKGVRKKLWLSIALIGSLQYGGLITSIVTQRK
55  KPFLSMMPVVAMYSLWLTKKYKKNVSYVCPNCHHVFNPSVIHFVTASHTPKTRKLQCPDC
HEMHYCVEIAKL*

Sequence 2545
Contig_0769_pos_850_251,
```

is similar to (with p-value 1.0e-51)
>sp:sp|P10585|GNTR_BACSU GLUCONATE OPERON TRANSCRIPTIONAL RE
PRESSOR (P28 PROTEIN). >pir:pir|C26190|C26190 transcription
1 repressor GntR - Bacillus subtilis >gp:gp|AB005554|AB00555
5   4_4 Bacillus subtilis genomic DNA, 36 kb region between gnt
and iol operons. NID: g2280496. >gp:gp|J02584|BACGNT_1 B.sub
tilis (gluconate operon) gntR, gntK and gntP genes encoding
gnt repressor, gluconate kinase and permease, and gntZ gene.
NID: g143013. >gp:gp|Z99124|BSUB0021_110 Bacillus subtilis
10   complete genome (section 21 of 21): from 3999281 to 4214814.
NID: g2636442.
gtgagtggtgagattgaatcccatatacaactaagtgaaaatcaagttgcacaacaattt
aatgttagtcgttcaccagtcagagatgcatttaagttacttcaaacagatcaactgatt
caattagaacgtatgggcgctcaggtacttccttttggtgatcaagagaaaaaggaaatg
15   tatgatttgcgtttgatgctcgaatcattcgcctttttcaaaattaagtggaacagataca
caacatattattaaggaaatgaaaaagcaattagaaatgatgaaggttgcagtccaattt
gaggatgctgaagcatttacacaacatgattttgagtttcatgaggtgatgattcaagca
acaaaccatcagtatcttaaagtgttttggaatcaccttaaacctgttatggaatcactc
atactcatttcaatgagacaaagaatggcaaatgaccccaaagatttcgagagaattcat
20   aaaaatcatcaagttttttatagatgctgttgagaacgatgatgcctccatattgagaaaa
gcattccatttaaattttgatgatgtaggagaaaatattgaagcattctggttacgttaa

Sequence 2546
25   VSGEIESHIQLSENQVAQQFNVSRSPVRDAFKLLQTDQLIQLERMGAQVLPFGDQEKKEM
YDLRLMLESFAFSKLSGTDTQHIIKEMKKQLEMMKVAVQFEDAEAFTQHDFEFHEVMIQA
TNHQYLKVFWNHLKPVMESLILISMRQRMANDPKDFERIHKNHQVFIDAVENDDASILRK
AFHLNFDDVGENIEAFWLR*

30   Sequence 2547
Contig_0770_pos_3341_3688,
is similar to (with p-value 3.0e-41)
>sp:sp|P37527|YAAD_BACSU 31.6 KD GUANYLYLATED PROTEIN IN DAC
A-SERS INTERGENIC REGION. >gp:gp|D26185|BAC180K_75 B. subtil
35   is DNA, 180 kilobase region of replication origin. NID: g467
326. >gp:gp|Z99104|BSUB0001_11 Bacillus subtilis complete ge
nome (section 1 of 21): from 1 to 213080. NID: g2632267.
atgtctaaaatagtaggatcagatcgagttaaaagaggaatggctgaaatgcaaaaaggc
ggtgtcattatggacgtcgttaatgcagaacaagctaaaattgctgaagaagccggagct
40   gttgccgtaatggcattagagcgtgtaccatcagatattcgtgctgctggcggtgttgca
cgtatggcgaatcctaaaatagttgaagaagttatgaatgccgtatcaattccggttatg
gctaaagccagaattggtcatattacagaagctagagttttagaatcgatgggtacacgg
tttcaagaacccaacatactacaaacgaatttcaaaaggcgagagtaa

45   Sequence 2548
MSKIVGSDRVKRGMAEMQKGGVIMDVVNAEQAKIAEEAGAVAVMALERVPSDIRAAGGVA
RMANPKIVEEVMNAVSIPVMAKARIGHITEARVLESMGTRFQEPNILQTNFKRRE*

Sequence 2549
50   Contig_0770_pos_5637_5939,
is similar to (with p-value 3.0e-42)
>gp:gp|AF016634|AF016634_1 Lactococcus lactis cremoris ClpB
chaperone homolog (clpB) and phosphoribosylformylglycinamide
cyclo-ligase (pur5) genes, complete cds; and phosphoribosyl
55   glycinamide formyltransferase (pur3) gene, partial cds. NID:
g3150045.
atgattcgaccttcagattctttcacttctttcaatactgctttttaaacgttcttcaaat
tcacctctatattttgcacctgcaactaaggcacttaaatcaagctcgaaaatcgtttta
tcgagtaatgattctggaacgtctttacgtacaattcgttgtgctaaaccttcaacaatt

624

gcagctttacctacacctggttcaccgattaaaaccggattattttttgttttccgactt
aatatacgaattgtattacgtatttcttcatctctaccgatgacagggtccattttacct
tga

Sequence 2550
MIRPSDSFTSFNTAFKRSSNSPLYFAPATKALKSSSKIVLSSNDSGTSLRTIRCAKPSTI
AVLPTPGSPIKTGLFFVFRLNIRIVLRISSSLPMTGSILP*

Sequence 2551
Contig_0770_pos_4797_3853,
is similar to (with p-value 0.0e+00)
>gp:gp|AF016634|AF016634_1 Lactococcus lactis cremoris ClpB
chaperone homolog (clpB) and phosphoribosylformylglycinamide
cyclo-ligase (pur5) genes, complete cds; and phosphoribosyl
glycinamide formyltransferase (pur3) gene, partial cds. NID:
g3150045.
gtggaaacagaaagagaaaagctattaagtttaagtgacatcttacacaaacgtgtagta
ggtcaagataaagcagttgatttagtatcagacgcagtagttagagcacgtgctggaatt
aaagatccgaatagaccaatcggaagtttcttattcttaggacctactggagtaggtaaa
actgaattagcaaaatcgcttgcttcatcacttttcgattctgaaaaacatatgattaga
attgatatgagcgaatatatggaaaaacatgctgtatcacgtttaattggtgcacctcca
ggttatgtaggtcacgatgaaggtggtcaattaactgaagcagttagacgtaatccatac
tcagttatttttgttagacgaagttgaaaaagcacatagcgatgtttttaatgtattactt
caaatactagatgaaggtcgtcttacggattctaaaggtagaagtgttgactttaaaaat
accattatcatcatgactagtaatattggttcacaagtattacttgaaaatgtaaaagat
gctggtgaaattagtgatgatacagagaaagcagttatggacagtctacatgcatacttc
aaacctgaaatattaaatcgtatggatgacatcgtgttatttaaaccattatcagttaat
gatatgagtatgattgtagataaaattttaacacaattaaatatgagattattagatcaa
catatctcaattgaagtgacagaagaagcgaaaaaatggctaggtgaagaagcgtatgaa
ccacaatttggtgcaagaccattaaaaacgctttgttcaacgacaaatagaaactccaatt
gcacgtatgatgattaaagaaagtctacctgaaggtacaataattaaagtagatttaaat
gacaataaagaacttgattttaaggttgttaaacctacgtcttaa

Sequence 2552
VETEREKLLSLSDILHKRVVGQDKAVDLVSDAVVRARAGIKDPNRPIGSFLFLGPTGVGK
TELAKSLASSLFDSEKHMIRIDMSEYMEKHAVSRLIGAPPGYVGHDEGGQLTEAVRRNPY
SVILLDEVEKAHSDVFNVLLQILDEGRLTDSKGRSVDFKNTIIIMTSNIGSQVLLENVKD
AGEISDDTEKAVMDSLHAYFKPEILNRMDDIVLFKPLSVNDMSMIVDKILTQLNMRLLDQ
HISIEVTEEAKKWLGEEAYEPQFGARPLKRFVQRQIETPIARMMIKESLPEGTIIKVDLN
DNKELDFKVVKPTS*

Sequence 2553
Contig_0770_pos_1762_383,
is similar to (with p-value 0.0e+00)
>sp:sp|P39616|DHA2_BACSU PROBABLE ALDEHYDE DEHYDROGENASE YWD
H (EC 1.2.1.3). >pir:pir|S39713|S39713 hypothetical protein
- Bacillus subtilis >gp:gp|X73124|BSGENR_59 B.subtilis genom
ic region (325 to 333). NID: g413923. >gp:gp|Z99123|BSUB0020
_91 Bacillus subtilis complete genome (section 20 of 21): fr
om 3798401 to 4010550. NID: g2636240.
atgacaataattagagataaatttaacaatagtaaagctttttttaatacgcataaaaca
aaaaaccttaaatttcgaaaacaacaacttaaattactaagtaaaaatatcaaaaatcat
gaaaatgaattattagatgccttatataaagatttaggtaaaagtaaggttgaagcatac
gcaactgaaattggtatgctttttgaaaagcataaagctaatgcgcaaagagttaaaaaat
tggtcgaaaaccaaacaaacggatacaccactctacttattccctacaaagagttatatt
aaaaaagaaccttacggtacggtgcttattataggaccatttaattatccggttcaatta
gttttcgagcctctcatcggagcaatagctgccggaaatactgctatagttaaaaccttca
gagttaacacctcatgttgccattgtgatcaaggacatcattgaagatacatttgatgaa
gcatacgtttctgttgttgaaggtggtattgaagaaacccaaacgttattaagtctàcca

```
      tttgattatatgttctttactggcagtgaaaaagtcggaaaaattgtctatgaagctgca
      gcaagaaaattaattccagttactcttgaacttggcggtaaatcacctgtcattgtcgat
      gatacagccaatatcaaagtagcgagtgaacgtattagtttttggtaaatttactaatgct
      ggtcaaacatgtgtcgctccagattatatattagttcagcggaaagttaaaaatgattta
   5  ataaaagctcttaaaaaaacaattactgaattttacggagaaaatattgaaaaaagccct
      gatttcggacggattgttaatcaaaaacactttaatcggttgaatgacttgattcaaatt
      cataaagataatgttgtttttggaggtaatagttctaaagaagatttatatattgaacct
      actttattggataacataaccaatgacaataaatcatgaaagaagaaatattcggtccc
      attttgcctattattacttatgataatttcgatgaagtacttgaaatcatccaaagtaaa
  10  tcaaaaccactaagtttgtatcttttttagcgaagatgaaaacatgacacatagagtggtt
      gaagaattatcatttgggggcggtgcaattaacgatacgttaatgcatttagctaatcct
      aacttacctttcggtggtgtaggttcttcaggcataggtcaatatcatggtaagtattct
      tttgatacatttagtcatatgaaatcatacacatttaaatctacacgtctagaatcgagt
      ttattttttccctccatataaaggtaaatttaaatatattaaaaccttcttcaagaactag
  15
```

Sequence 2554
MTIIRDKFNNSKAFFNTHKTKNLKFRKQQLKLLSKNIKNHENELLDALYKDLGKSKVEAY
ATEIGMLLKSIKLMRKELKNWSKTKQTDTPLYLFPTKSYIKKEPYGTVLIIGPFNYPVQL
VFEPLIGAIAAGNTAIVKPSELTPHVAIVIKDIIEDTFDEAYVSVVEGGIEETQTLLSLP
FDYMFFTGSEKVGKIVYEAAARKLIPVTLELGGKSPVIVDDTANIKVASERISFGKFTNA
GQTCVAPDYILVQRKVKNDLIKALKKTITEFYGENIEKSPDFGRIVNQKHFNRLNDLIQI
HKDNVVFGGNSSKEDLYIEPTLLDNITNDNKIMKEEIFGPILPIITYDNFDEVLEIIQSK
SKPLSLYLFSEDENMTHRVVEELSFGGGAINDTLMHLANPNLPFGGVGSSGIGQYHGKYS
FDTFSHMKSYTFKSTRLESSLFFPPYKGKFKYIKTFFKN*

Sequence 2555
Contig_0771_pos_3650_4303,
is similar to (with p-value 7.0e-69)
>gp:gp|AF008219|AF008219_3 Borrelia afzelii R-IP3 chromosome
   right end, arcA and arcB genes, complete cds. NID: g2697111

```
      atgaaaaatttacgtaacagaagcttttttaactttattagacttttcacgacaagaggta
      gaattttttattaacactctccgaagatttgaagcgtgccaaatatatcggcactgaaaag
      cctatgctaaaaaataaaaatatcgcgcttctttttgaaaaagattccactagaacacgt
      tgcgcattcgaagttgccgcacatgatcaaggtgcacacgtcacttatcttggacctaca
      ggttctcaaatgggtaaaaaagaaactgctaaagatacagcacgtgtacttggtggtatg
      tatgatggtattgagtaccgaggtttctctcaacgtactgtagaaacattagcgcaatat
      tcaggtgttccggtatggaatggattaaccgatgaagatcaccctacacaagtgcttgct
  40  gacttttaactgctaaagaagtattgaaaaaagagtatgctgatatcaactttacttat
      gttggcgatggacgtaacaatgttgctaacgcattaatgcaaggtgctgccattatgggt
      atgaatttccatcttgtttgtcctaaagaactcaatccgacagaagaattattaaatcgt
      tgcgaacgtattgcgacggaaaatggcggtaacattttaatatcattactttaa
```

Sequence 2556
MKNLRNRSFLTLLDFSRQEVEFLLTLSEDLKRAKYIGTEKPMLKNKNIALLFEKDSTRTR
CAFEVAAHDQGAHVTYLGPTGSQMGKKETAKDTARVLGGMYDGIEYRGFSQRTVETLAQY
SGVPVWNGLTDEDHPTQVLADFLTAKEVLKKEYADINFTYVGDGRNNVANALMQGAAIMG
MNFHLVCPKELNPTEELLNRCERIATENGGNILISLL*

Sequence 2557
Contig_0771_pos_5383_4310,
putative peptide of unknown function
```
      atgattaattcatttggaatggttgtaggtaatttgttaggtggtttttttgtttgataaa
  55  cttggtggttacaaaacgattttaataggtacatttacgtgtttatgtagtaccacatta
      ctcaacttgtttcatggctggccatggtatgcaatttggttagtactacttggatttggt
      ggcggaatgatagttcctgctatttatgcgatggcaggtgccgtttggcctaatggagga
      agacaaacttttaatgctatctacttagcacaaaatataggggttgctttaggcgcggca
      ttaggaggttttgtggctgaatttagtttcaattatattttttatggcaaacctcattatg
```

```
tatgtactttttgccatcgtggcgattacacagttcaatttagagattaatgctaaattt
aaaccacaagattcgatagatttaaaaagcaaagaaaataaaaaacgatttactgctatg
atgctagtatgtgcaatgtttgcaatttgttggattgcatatattcaatgggaaacaacg
atagcttcattcacacaatcaattaatatttcaatgtctcaatatagtgtattatcgaca
attaatggaattatgattttagtagctcaacctttaataagaccaattattatcttatta
aaaggtaatttaaaaacatcaaatgtttgtaggtattttaatttttatgagttctttccta
gtgacgagttttgcaaatcactttgctatatttgtagttggcatggtcattttaacttttt
ggagaaatgtttgtttggcctgcagtaccaactatagcaaatcaacttgcaccagttgga
aagcaagggcaataccaaggatttgttaattcagcatctacagtgggtaaagcatttgga
ccatttattggggggtatacttgtagatacattaatatgagtatgatgtttattgggatg
attatattattaagttttgcactgttattttttaagtttctatgataaagtgttacccaag
aatttaaaaatcaacatcaatcaagaagacgacgaaatcagaatggtatttaa
```

Sequence 2558
```
MINSFGMVVGNLLGGFLFDKLGGYKTILIGTFTCLCSTTLLNLFHGWPWYAIWLVLLGFG
GGMIVPAIYAMAGAVWPNGGRQTFNAIYLAQNIGVALGAALGGFVAEFSFNYIFMANLIM
YVLFAIVAITQFNLEINAKFKPQDSIDLKSKENKKRFTAMMLVCAMFAICWIAYIQWETT
IASFTQSINISMSQYSVLWTINGIMILVAQPLIRPIIILLKGNLKHQMFVGILIFMSSFL
VTSFANHFAIFVVGMVILTFGEMFVWPAVPTIANQLAPVGKQGQYQGFVNSASTVGKAFG
PFIGGILVDTFNMSMMFIGMIILLSFALLFLSFYDKVLPKNFKNQHQSRRRRNQNGI*
```

Sequence 2559
Contig_0771_pos_2844_2416,
is similar to (with p-value 2.0e-50)
>pir:pir|S58181|S58181 fofB protein - Staphylococcus sp. >gp
:gp|X89875|SSPPDNAFB_1 Staphylococcus sp. plasmidic DNA for
fosB gene. NID: g927563.
```
atggaaataacaaatgttaatcatatttgttttttcagtgagtgatttaaatacctctata
caattttataaagatattttacatggtgacttattagtatcagatagaacgacagcatat
ttaactattggtcatacttggattgcactgaatctagaaaaaaatataccaaggaatgaa
ataagtcattcctatacgcacgttgcttctccatagatgaagaagattttcaacagtgg
attcaatggcttaaagagaatcaagtaaatattttaaaagggcgaccaagagacattaaa
gacaaaaaatcgatatattttacagatctggatgggcataaaattgaattacatactgga
acattaaaagatagaatggaatattataaatgtgagaagacgcatatgcaattttacgat
gagttttga
```

Sequence 2560
```
MEITNVNHICFSVSDLNTSIQFYKDILHGDLLVSDRTTAYLTIGHTWIALNLEKNIPRNE
ISHSYTHVAFSIDEEDFQQWIQWLKENQVNILKGRPRDIKDKKSIYFTDLDGHKIELHTG
TLKDRMEYYKCEKTHMQFYDEF*
```

Sequence 2561
Contig_0771_pos_1176_178,
is similar to (with p-value 0.0e+00)
>sp:sp|P53557|BIOB_BACSU BIOTIN SYNTHETASE (EC 2.8.1.-). >gp
:gp|AF008220|AF008220_77 Bacillus subtilis rrnB-dnaB genomic
 region. NID: g2293135. >gp:gp|U51868|BSU51868_5 Bacillus su
btilis biotin biosynthetic operon genes, complete and partia
l cds. NID: g1277024. >gp:gp|Z99119|BSUB0016_93 Bacillus sub
tilis complete genome (section 16 of 21): from 2997771 to 32
13410. NID: g2635411.
```
atgctaattttttaagaaaaaggagttaaagattatgacattaaacctagctcaacgtgtg
ttaaatcaagagtcattaacaaaagatgaagcaatatctattttcgaaaatgctgaaatt
gatacatttgatttattaaatgaagcctacacagtgagaaaacattactatggtaaaaaa
gttaagcttaatatgatattaaatgctaaaagtggtatctgtgcagaaaattgtgggtac
tgtgggcaatctgtaaaaatgaaagaaaagcaacgttatgcacttgttgaacaggaccaa
attaaagaaggcgctcaagtggcaactgaaatcaaatcggtacatactgtattgttatg
agtggtagaggtcctagtaacagagaagtcgatcatatttgcgaaacagtagaagatatt
aaaaagatacacccacaactaaagatttgtgcgtgcttaggattaacgaaagaagaacag
```
```

627

```
gctaaaaaattaaaggctgctggtgtcgatcgttataatcataatttaaatacgagtgag
cgttatcacgatgaagtagtaactacacatacatatgaggatagagtgaatacggttgaa
atgatgaaagataataatatttctccttgttcaggtgtgatatgtggtatgggagagtcg
aatgaggacattattgatatggcatttgctttaagagccatcgatgctgatagcattcct
attaattttttacatcctattaaaggaactaaatttggtggattagatttattgtcacca
atgaaatgtttaagaattatagcgatgtttaggttaatcaatccaacaaaagaaattcga
attgcaggtggacgggaggtaaatctacgttcattacaaccactcgcattgaaagcggct
aattcaatttttgtaggagattacttaattacaggcggtcaaccgaatgaggaagattat
cgcatgattgaagatttagggtttgaaatcgacagttaa
```

Sequence 2562
MLIFKKKELKIMTLNLAQRVLNQESLTKDEAISIFENAEIDTFDLLNEAYTVRKHYYGKK
VKLNMILNAKSGICAENCGYCGQSVKMKEKQRYALVEQDQIKEGAQVATENQIGTYCIVM
SGRGPSNREVDHICETVEDIKKIHPQLKICACLGLTKEEQAKKLKAAGVDRYNHNLNTSE
RYHDEVVTTHTYEDRVNTVEMMKDNNISPCSGVICGMGESNEDIIDMAFALRAIDADSIP
INFLHPIKGTKFGGLDLLSPMKCLRIIAMFRLINPTKEIRIAGGREVNLRSLQPLALKAA
NSIFVGDYLITGGQPNEEDYRMIEDLGFEIDS*

Sequence 2563
Contig_0774_pos_436_873,
is similar to (with p-value 3.0e-34)
>sp:sp|P39149|UPP_BACSU URACIL PHOSPHORIBOSYLTRANSFERASE (EC
  2.4.2.9) (UMP PYROPHOSPHORYLASE) (UPRTASE). >pir:pir|S49364
  |S49364 uracil phosphoribosyltransferase - Bacillus subtilis
  >gp:gp|Z38002|BSSPORUPP_10 B.subtilis spoII-R, glyC and upp
  genes. NID: g556877. >gp:gp|Z99122|BSUB0019_186 Bacillus su
  btilis complete genome (section 19 of 21): from 3597091 to 3
  809700. NID: g2636029.
atgttaaaagtattcaacactgaatcaagcatctctctatgttgggcggtgactacaaca
acgggttccaggtcagaatctttctctaacgtttaatcaacggagccattttatagct
tcaggcctagttccaaatatggtcataactttttcatcaaaaatatctatctccgaatc
tattatttagtaccgaataatctatcccctgcatcacctaaacctggtgtaatatacgct
ttgtcatttaattttcatctaatgccgcaatatatatctacatctgggtgtgcttct
tgcatttttttcaacgccttcaggggcagctattaaacacataaaacgtatacttttagct
ccacgttttttttaatgaagaaattgcttcaatagctgaagcaccagtagcaagcatagga
tcaaccacaataatttga

Sequence 2564
MLKVFNTESSISLCWAVTTTTGSRSESFSNVLINGAIFIASGLVPNMVITFFIKNIYLRI
YYLVPNNLSPASPKPGVIYALSFNFSSNAAIYISTSGCASCIFSTPSGAAIKHIKRILLA
PRFFNEEIASIAEAPVASIGSTTII*

Sequence 2565
Contig_0774_pos_2423_3718,
is similar to (with p-value 9.0e-29)
>pir:pir|S57509|S57509 integral membrane protein - Streptomy
ces pristinaespiralis >gp:gp|X84072|SPDNAPTR_1 S.pristinaesp
iralis ptr gene. NID: g872305.
gtgccattacaatcatcatacaatagtgatattggtactattaatatagcagttagctta
tcggcactattttctggtctgtttattgtaggtgcaggagatattgcagataaaattggt
agagtaaaagtgacgtacataggcttagcacttaatattgttgggtcgattttaatcatt
attacgccattaccaagtctattgattattggacgtgctattcaaggattgtcagcggca
tgtataatgccagcgacactcgcaatcattaatgaatattatatcgggacagcacgacaa
cgtgcattaagctactggtccatcggttcatggggaggtagtggtgtttgtactttgttt
ggtggtttaatggcaactaaccttggatggcgctcaatctttattgtttcaattatctg
acaatattatccatgtttctcattaaacatacacccgaaacaaaagcagagcctatagga
gatcaaccgacagagacaaagaaatttgatgttgttggtttaatcatcttagtggttagt
atgttaagtattaatgtgataataactcaaacctctcaatttggtttgttctcaccattt
attttgggacttattgcaattttttgttatatcgttaattatattcgtgatttacgaaaat
```

```
      aaaatcaaacaaccacttgtcgattttgatatctttaaaaacaaaggttatacaggtgca
      acgatttcaaactttatgttgaatggtgtagctggtggtacattgattgtagtaaatact
      ttttatcagcaaaaattagattttaactctcaggaaacaggatatatttcacttacatac
      ctaattgcagtattaattatgatacgtgtgggtgaaatgatattacaatcgttaggacct
  5   aaaagacctttgttactcggaagtgccttgaccgtcataggattaatattattatctttg
      acgtttttacctaatgcttggtatatagcgtcaagtgtcattggttatttattatttggt
      accggtttaggtgtttatgcaacaccatccacggatacagctgttgcacaagcacragat
      gataaagtaggcgtcgcatccggtgtatataagatggcatcatccttaggaaatgcattt
      ggtgtggccatctcaagtacagtttacagcgtacttgcagcccaacttaatctgacttta
  10  ggtggtttttactggagtaatgtttaatgcgcttatagcattattagcattcctttctatt
      ttgttcttaataccgaaaaaacagtctaatgtataa
```

Sequence 2566
VPLQSSYNSDIGTINIAVSLSALFSGLFIVGAGDIADKIGRVKVTYIGLALNIVGSILII
ITPLPSLLIIGRAIQGLSAACIMPATLAIINEYYIGTARQRALSYWSIGSWGGSGVCTLF
GGLMATNLGWRSIFIVSIILTILSMFLIKHTPETKAEPIGDQPTETKKFDVVGLIILVVS
MLSINVIITQTSQFGLFSPFILGLIAIFVISLIIFVIYENKIKQPLVDFDIFKNKGYTGA
TISNFMLNGVAGGTLIVVNTFYQQKLDFNSQETGYISLTYLIAVLIMIRVGEMILQSLGP
KRPLLLGSALTVIGLILLSLTFLPNAWYIASSVIGYLLFGTGLGVYATPSTDTAVAQAPD
DKVGVASGVYKMASSLGNAFGVAISSTVYSVLAAQLNLTLGGFTGVMFNALIALLAFLSI
LFLIPKKQSNV*

Sequence 2567
Contig_0774_pos_4133_5074,
putative peptide of unknown function
```
      atggaatatatgaaaatagcaattgcaggttctggcgcattaggtagtggatttggitgct
      aagttgtttcaacatggttatgacgtcactttaattgataattgggaacctcaagttact
      acaatacaacaggacggtctacatatcgatattaatggtgaagcgcatcatttcaggcta
      cctatgtatagactaacggaaattcctaaagcaacgtcctatgatattgttttctattt
      cctaaatctatgcaattagaagaggtgcttagtcatattcaacccatcttcatgataat
      acaattgttgtgtgcactatgaatggtttgaaacatgaacgtcttatacaacaatatgtt
      tctatagatagaattgtacgtggagtaacaacgtggactgccggtattgatcaacctggt
      cacacgcacttaatggggcaaggtcctgttgaaattgggtgtctcgtgcccgagggaaaa
      gaaagcgtagatatcattgttaatctgctacaaaatgcagaattaaaaggtgtaaaaagt
  35  gaacatttacatcaatcaatttggaagaaaatatgtgttaatggaacagctaattcatta
      tgtactatacttgaatgtaatttggcagcactgaataatagtgatgacgctaaaaatttg
      atatataaaattacacaagaaattgttcatgttgcaacagttgatgatgttcatcttaat
      gttgatgagatttttgattacttaattgctttaaatgataaagtaggcccacactatcct
      tctatgtaccaagacttaattaaagataatcgaacaactgaaatagattatattaatgga
  40  gcagttagtaaattagggaaagagaatcatattgctacacctgtaaatgattttgtaaca
      aatcttgtacatgctaaagaaaatcaacgtggtgcacaatga
```

Sequence 2568
MEYMKIAIAGSGALGSGFGAKLFQHGYDVTLIDNWEPQVTTIQQDGLHIDINGEAHHFRL
PMYRLTEIPKATSYDIVFLFPKSMQLEEVLSHIQPHLHDNTIVVCTMNGLKHERLIQQYV
SIDRIVRGVTTWTAGIDQPGHTHLMGQGPVEIGCLVPEGKESVDIIVNLLQNAELKGVKS
EHLHQSIWKKICVNGTANSLCTILECNLAALNNSDDAKNLIYKITQEIVHVATVDDVHLN
VDEIFDYLIALNDKVGPHYPSMYQDLIKDNRTTEIDYINGAVSKLGKENHIATPVNDFVT
NLVHAKENQRGAQ*

Sequence 2569
Contig_0774_pos_2605_2297,
putative peptide of unknown function
```
      atgattaaaatcgacccaacaatattaagtgctaagcctatgtacgtcactttтactcta
      ccaatttтatctgcaatatctcctgcacctacaataaacagaccagaaatagtgccgat
      aagctaactgctatattaatagtaccaatatcactattgtatgatgattgtaatggcacg
      actagatttacaagtgattgtgcaaataaccaaaatgtaataaccccctaaaataatccct
      aataataagcgattatctcccctaaacttttgtgatgtattcatctatcacttactcctt
      cacaattaa
```

Sequence 2570
MIKIDPTILSAKPMYVTFTLPILSAISPAPTINRPENSADKLTAILIVPISLLYDDCNGT
TRFTSDCANNQNVITPKIIPNNKRLSPLNFCDVFIYHLLLHN*

Sequence 2571
Contig_0774_pos_0_597,
is similar to (with p-value 3.0e-65)
>gp:gp|U81973|SAU81973_16 Staphylococcus aureus capsule gene
  cluster Cap5A through Cap5P genes, complete cds. NID: g1773
339.
atgaaaaaagttatgaccatatttggaactaggcctgaagctataaaaatggctccgttg
attaaaacgttagagaaagattctgacctggaacccgttgttgtagtcaccgcccaacat
agagagatgcttgattcagtgttgaatacttttaacataagtgcagattatgatttgaat
attatgaaagctggtcaaacattgtctgaagtaacatctgaagcaatgaaaaagttagaa
gatatcatacaaaaggaagtgcctgatatggtacttgttcatggtgatacagtgacaacc
tttctggagcattagccgcattttatagtcaaacacctataggacatgttgaagctgga
ttaagyagttataataaatattcaccttatcctgaagaaataaatagacaaatggttggg
gtaatggcagatttgcactttgccccaacctataatgctgcacagaatttagtaaaagag
ggtaaattagccaaacatatagctatcactggtaatacagctattgacgcaatgaattat
acaatcgatcaccaatattcatcatctatcatacaaaaacataaaaataaaaaTACT

Sequence 2572
MKKVMTIFGTRPEAIKMAPLIKTLEKDSDLEPVVVVTAQHREMLDSVLNTFNISADYDLN
IMKAGQTLSEVTSEAMKKLEDIIQKEVPDMVLVHGDTVTTFSGALAAFYSQTPIGHVEAG
LRSYNKYSPYPEEINRQMVGVMADLHFAPTYNAAQNLVKEGKLAKHIAITGNTAIDAMNY
TIDHQYSSSIIQKHKNKNT

Sequence 2573
Contig_0775_pos_5156_0,
putative peptide of unknown function
gtgtttatagctgctgaacatcaaacgtcagttgatgagttagagtatatttatgttcca
cctttaactacagaagctcaaaatgcctttgacgaacttaagcatgtgatgccaaaagta
tgttataaagccaatttagaacgtttggcatcgttgccaaatataaataacactgaccat
aatcctaatgctgaagcgcatcgtcacgctagcgattggagtgaagttcgtccagaatgg
ggtctagcacgaaatgctgaattcattattgggaaacgtcaaatcacccaaaatagtaat
ctagaggacgggcatttcttcataattatgattggacaaaggatgaagacggtgagatt
ttaactacaattatttctgggccagcactagtagcacaatggattaatttacaatactac
gcctcaaccgtggcacctcactattatggaagcggtagtaaaacaacgcaaactgtaaca
agtggtgtaggtgtcatgcaaggaaatgctagtgatttaatgtatggcttaccatggcag
tcagtaatgatgaatgacaaagaggcgtatcacgcacctattaggcttttaattgttatt
caagcgccagatgcatatattcaacgtttgttaaaacatcataatcactttagacaaaag
gttgatcatcaatggataagacttgccagtattgatgaaataatagttggaaagactgg
ta

Sequence 2574
VFIAAEHQTSVDELEYIYVPPLTTEAQNAFDELKHVMPKVCYKANLERLASLPNINNTDH
NPNAEAHRHASDWSEVRPEWGLARNAEFIIGKRQITQNSNLEGRAFLHNYDWTKDEDGEI
LNTIISGPALVAQWINLQYYASTVAPHYYGSGSKTTQTVTSGVGVMQGNASDLMYGLPWQ
SVMMNDKEAYHAPIRLLIVIQAPDAYIQRLLKHHNHFRQKVDHQWIRLASIDENNSWKDW
·X

Sequence 2575
Contig_0775_pos_3674_2475,
is similar to (with p-value 2.0e-92)
>sp:sp|P44953|DFP_HAEIN DNA/PANTOTHENATE METABOLISM FLAVOPRO
TEIN HOMOLOG. >pir:pir|G64104|G64104 pantothenate metabolism
  flavoprotein (dfp) homolog - Haemophilus influenzae (strain
  Rd KW20) >gp:gp|U32776|U32776_9 Haemophilus influenzae Rd s

ection 91 of 163 of the complete genome. NID: g1573969.
atgaaacatattttattagctgttacaggcggtatcgcagcatataaagcaattgattta
acaagtaaattaatacaatccggctatgatgtaagagttatgctatctgatcatgctcaa
gagtttgttactccgctagcttttcaagcaatcagtagaaatcctgtttacacaaataca

5   tttaaagaagaaatcctgaagagattcaacatgtatcattaggagactgggcagatgcg
attatagtcgcgccagcaactgctaatactatcgcaaaattaagtgttggaattgctgat
gatttaattacttctacattacttgctacaacaacaccaaaattcgttgcacccgcaatg
aatgtaaatatgtataacaatccacgtactaaacataatatgaaagtgctaagtcaagac
ggatattattttattgaacctggtagtggctatttagcatgtggttatgtagcaaaaggg

10   cgaatggaagaacccatgcaaatcctatctgttattaataaattttttactcaacagaag
aatgttgtcaaaagctcttttctggaaagcgcgcattagttacagctgggcctacagtt
gaagttattgatcctgttcgatacgtatcaaatcgttcatcaggaaaaatgggatatgct
atagctgaagcattacgagataagggagcaatcgtaactttaattagtggtcccacccac
ttatctctacctgaagggattaatgtagtaaaagttgagagtgcagatgatatgtttcaa

15   gctgtaaccgaacgctttgcgaaacaagatatagtgattaaagcagcggcggtgtctgat
tatacaccaatggacatacttgaacataaattaaaaaaacaagaaggaggattatctgtt
caatttaagcgtacaaaggatattttaaaatacttaggagaaaataaaacgcaccaatat
cttgttggttttgctgctgaaacacaaaatattgaacagtatgctctagacaaactcaaa
agaaaaaatgcagatgttatcatttcgaacaatgtaggtgatacatccataggctttagt

20   tcagatgacaatgaattaactatgcatttttaaaaataatgaaaaagtaaatattaagaaa
ggaaaaaaatcagctttagcacatcaaattatagaaattttagaaactaggtggcagtaa

Sequence 2576
25   MKHILLAVTGGIAAYKAIDLTSKLIQSGYDVRVMLSDHAQEFVTPLAFQAISRNPVYTNT
FKEENPEEIQHVSLGDWADAIIVAPATANTIAKLSVGIADDLITSTLLATTTPKFVAPAM
NVNMYNNPRTKHNMKVLSQDGYYFIEPGSGYLACGYVAKGRMEEPMQILSVINKFFTQQK
NVVKSSFSGKRALVTAGPTVEVIDPVRYVSNRSSGKMGYAIAEALRDKGAIVTLISGPTH
LSLPEGJIIVVKVESADDMFQAVTERFAKQDIVIKAAAVSDYTPMDILEHKLKKQEGSLSV

30   QFKRTKDILKYLGENKTHQYLVGFAAETQNIEQYALDKLKRKNADVIISNNVGDTSIGFS
SDDNELTMHFKNNEKVNIKKGKKSALAHQIIEILETRWQ*

Sequence 2577
Contig_0775_pos_2367_67,
35   is similar to (with p-value 0.0e+00)
>sp:sp|P94461|PRIA_BACSU PRIMOSOMAL PROTEIN N' (REPLICATION
FACTOR Y) (FRAGMENT). >gp:gp|Y10304|BSPRIADFS_1 B.subtilis p
riA, def, fmt, sun genes. NID: g1772497.
gtggtaccttttggccctagaactatacaaggatacgtcatgaatattcaacaaaaacca

40   gatggaaatatggatatatcgaaactaaaagaaataaaagaagtacgtgatattaaacct
gaattaacatccgaactgattcaattaagcgaatggatgagccattatcatgtgatgaaa
cgtatttctgttttagaagcgatgttgccaagtgccattaaagcaaagtataagaaagct
ttttcaattatcgatccaaaaaatttatcttcaaaaaccaaagcgctatttaacaatgac
ggttattacttatataaagaagttcagcaaacaatgatttagaagaaatgttgactttg

45   ttaaatcaaggattgattgaagaggtcacgatactttctcaaaacacaaaaaagaaaact
caaaaagctgttggcgtagttaatacgttgaatggtgatgaagtacttgcaaaactcgag
aaatatacaaaacaatatgatttgtatgcatttttattagaagagtctcatcgaacagtg
tttttaaaagaaatcaatgatatgggcttctctcactcgagtttagattctttaatcaaa
aaaggctatattgaaaaatatatcgccgaggttttcagagatccatatgcaaatcctata

50   tttgaacaagaacaaagaggatattaactaaagaacagcaagatgcatttgaagctatt
caacattatattcatgatgaaaaagaaagaacattttattacacggagtcacaggttca
ggtaaaaccgaagtctatcttcaaacaatagaagaagttcttaataaaggtaaagaagcc
atgatgttagtgccggaaatcgccttaacacctcagatggtactaagatttaaacgtcga
tttggagatgatgtagcggtattacattccggactttcaaaaggtgaacgttatgatgag

55   tggcaaaaaattagagacggtcgagctcgagtgagtgtaggtgctcgttcaagtattttc
gcaccgtttaaaaatttaggcatcattataattgatgaagaacatgaatctacatataaa
caagaagattatcccagatatcatgcacgtgatattgcacaatggagaagtcaatttcat
cattgtcctgtagtttaggtagcgcaacaccgagtcttgagtcatatgcaagagcagaa
aaaaatgtttacgagttgttgtcattgccacatagagtcaatcaacaagcgttaccgcat

atgata\.tatagatatgagagaagaattaagtgaaggtaatcgttccatgtttt\.tata
gcactaagacaagcgatacaagaacggttggataaaaaagaacaaatagtactattctta
aatagaaggggatatgcttcatttatgttatgtagagattgtggttacgttccccaatgt
ccccattgtgatatttcgttaacatatcataaaacaaccgatcaattaaaatgtcattac

5   tgtggttatcaagaaaatccaccatctcaatgtccaaattgtgaaggtgatcatatcaga
caagtcggaactggaacgcaacgtgtagaagaattattacaacaagaattccctcatgct
cgtattataaggatggatgttgatacaacttcaagaaaaggtgcacatgagaaattgtta
aatgactttgaagcaggaaaaggagatatcttattaggtacgcaaatgattgctaaaggt
ttggattatcctaacattactctagttggtgtgctcaatgctgacactatgttaaactta

10 cctgactttcgtgccagtgaacgaacataccaacttttaactcaggtatctggacgcgca
ggtcgtcatgaaaaagaaggacaagttatcatacaaacgtacaaccctgatcattattca
ataaaggatgtgaaattaaacgattatctttcttttttatcaaaaagaaatgaattatcga
aaattaggaaaaatatccaccttactattttttgattaactttaccatttcacatactgat
ataaaaaaggtcatgatggcctctaagcatatacatcaaatttttagtacagcacttaagt

15 gaaaaagcattcgtgctaggcccttcaccagcagcactagcaagaattaacaatgagtat
cgttttcaaatactagtaaaatataagagtgagcctcaattacatcaagcgttacaatat
ttagatgattataatcatgatcaatatgtaaaggataaactatcattaaaaattgatatc
aatccacaaatgatgatgtga

20 Sequence 2578
    VVPFGPRTIQGYVMNIQQKPDGNMDISKLKEIKEVRDIKPELTSELIQLSEWMSHYHVMK
    RISVLEAMLPSAIKAKYKKAFSIIDPKNLSSKTKALFNNDGYYLYKEVQQNNDLEEMLTL
    LNQGLIEEVTILSQNTKKKTQKAVGVVNTLNGDEVLAKLEKYTKQYDLYAFLLEESHRTV
    FLKEINDMGFSHSSLDSLIKKGYIEKYIAEVFRDPYANRIFEQEQKRILTKEQQDAFEAI
25 QHYIHDEKERTFLLHGVTGSGKTEVYLQTIEEVLNKGKEAMMLVPEIALTPQMVLRFKRR
    FGDDVAVLHSGLSKGERYDEWQKIRDGRARVSVGARSSIFAPFKNLGIIIIDEEHESTYK
    QEDYPRYHARDIAQWRSQFHHCPVVLGSATPSLESYARAEKNVYELLSLPHRVNQQALPH
    IDIIDMREELSEGNRSMFSIALRQAIQERLDKKEQIVLFLNRRGYASFMLCRDCGYVPQC
    PHCDISLTYHKTTDQLKCHYCGYQENPPSQCPNCEGDHIRQVGTGTQRVEELLQQEFPHA
30 RIIRMDVDTTSRKGAHEKLLNDFEAGKGDILLGTQMIAKGLDYPNITLVGVLNADTMLNL
    PDFRASERTYQLLTQVSGRAGRHEKEGQVIIQTYNPDHYSIKDVKLNDYLSFYQKEMNYR
    KLGKYPPYYFLINFTISHTDIKKVMMASKHIHQILVQHLSEKAFVLGPSPAALARINNEY
    RFQILVKYKSEPQLHQALQYLDDYNHDQYVKDKLSLKIDINPQMMM*

35 Sequence 2579
    Contig_0776_pos_1350_2054,
    is similar to (with p-value 2.0e-50)
    >sp:sp|P54163|YPDP_BACSU HYPOTHETICAL 25.7 KD PROTEIN IN BCS
    A-DEGR INTERGENIC REGION. >gp:gp|L77246|BACYACA_10 Bacillus
40 subtilis (YAC10-9 clone) DNA region between the serA and kdg
    loci. NID: gl256615. >gp:gp|Z99115|BSUB0012_140 Bacillus su
    btilis complete genome (section 12 of 21): from 2195541 to 2
    409220. NID: g2634478.
    atgtataatgaaatatttggtattgcgtcatttattgttacattcgctttaatggtactg
45 atgtatcgctgttttggtaaacaaggactaattgcttgggtagcaataggaacgattatc
    gctaatatacaggtcataaaagcggttcatattttggtattacggctacacttggaaat
    gtcatgtttgcttctatatatttagctactgatatattaaatgacatctatggtcgtaaa
    gttgctaaaagagcggtgtggcttggtttctcttctaccttagtaatgattatagtcatg
    caaatgtcattgcattttattcctgccccagtagacaatgcgcaaaactcattaaaaatg
50 atttttgatttagtgcctagaattgctataggttccattattgcttatatcataggccaa
    catattgatgtatttatattcagtatgattaaaaagatatttagctctgataagacctt
    tttattagagcatatggtagtaccattttaagttctatcattgataccggtttatttgtt
    tcaattgcttttattggtactatgcctggtactgctgtttttgaaatatttattaccact
    tacttgttaaaactagtgtcaactattttta'tgtaccatttggatatatcgctaagtca
55 ctatatcgaaaaggaaagatagaacaactagataatgggtattga

    Sequence 2580
    MYNEIFGIASFIVTFALMVLMYRCFGKQGLIAWVAIGTIIANIQVIKAVHIFGITATLGN
    VMFASIYLATDILNDIYGRKVAKRAVWLGFSSTLVMIIVMQMSLHFIPAPVDNAQNSLKM

IFDLVPRIAIGSIIAYIIGQHIDVFIFSMIKKIFSSDKTFFIRAYGSTILSSIIDTGLFV
SIAFIGTMPGTAVFEIFITTYLLKLVSTIFNVPFGYIAKSLYRKGKIEQLDNGY*

Sequence 2581
Contig_0776_pos_2379_2771,
is similar to (with p-value 3.0e-19)
>sp:sp|P36921|EBSB_ENTFA CELL WALL ENZYME EBSB. >pir:pir|B49
939|B49939 ebsB protein - Enterococcus faecalis >gp:gp|L2380
2|ENEEBSA_2 Enterococcus faecalis pore forming, cell wall en
zyme, regulatory, and dehydroquinase homologue proteins (ebs
A,ebsB,ebsC,and ebsD) genes, complete cds with repeat region
. NID: g388106.
atggctaaaatacattttgatgctgcgactaaaggaaatcccggccgaagtgcttgtgcg
attattattaaagaaaaattcacaaagatatacatttacccatgatttaggtgaaatggat
aatcatagtgcagaatgggcagcaatgttacacgctttggaacatgcacgcgaattaaaa
gtatctaacgcgttactttttactgattcaaaattaattgaagatagtatgatgcaaggt
aaagttaaaaatgctaagtttaaagtttattttgaaaacatagaaatcttagagcaaagt
tttgatttgatgtttgtgagatggattccacgaaagcaaaataaagaagcgaatcaactt
gctcaacaaacactatacaaacttacatcataa

Sequence 2582
MAKIHFDAATKGNPGRSACAIIIKENSQRYTFTHDLGEMDNHSAEWAAMLHALEHARELK
VSNALLFTDSKLIEDSMMQGKVKNAKFKVYFENIEILEQSFDLMFVRWIPRKQNKEANQL
AQQTLYKLTS*

Sequence 2583
Contig_0776_pos_5545_4955,
is similar to (with p-value 3.0e-22)
>gp:gp|AL034447|SC7A1_23 Streptomyces coelicolor cosmid 7A1.
NID: g4007715.
gtgttcgttatggaaaaagaagaaatctttcttcctgaaagtgaaaaaaatgggaaacct
atcgacttatttttttatatggtttgcagccaatctaggtattctcgggatagtttatggc
gctgtcatagtaagttacggattgagcttttacaatcaattttgattgctataatagga
ccgttatcctttgttctagttggttatataagtgtagctggtagagatagtggagctata
acttttatgctctcaagagcaccatttggatttaaaggcaatcatataacctgctttaatt
ggctgggtaggtcaagttggttggttatctgttaatgtttctacaggaactttaactctt
ctggctttattcaatacttttggttttaagactagtacatttctaattttgatgagttta
gcgattttttgctgggctagttattatatctgttctttttttcacaaaaagtacttgtatca
gtacaaacattttttcacatatgtatttggtgcattaaccttattagttataacaatttta
attactaatactgattggaacgccctttttttctatgaaatctgggtcttaa

Sequence 2584
VFVMEKEEIFLPESEKNGKPIDLFFIWFAANLGILGIVYGAVIVSYGLSFLQSILIAIIG
PLSFVLVGYISVAGRDSGAITFMLSRAPFGFKGNHIPALIGWVGQVGWLSVNVSTGTLTL
LALFNTFGFKTSTFLILMSLAIFAGLVIISVLFSQKVLVSVQTFFTYVFGALTLLVITIL
ITNTDWNALFSMKSGS*

Sequence 2585
Contig_0776_pos_4784_2949,
putative peptide of unknown function
gtgataaaagagttatcacaaaagaaacgagatgcaataaataacaacactgatttaaca
ccttctcaaaggcacatgctttagcagatattgataaaacagaaaaagatgcactcaa
catatcgaaaattctaattcaattgatgatatcaataacaataaagagcatgcatttaat
actttagctcatatcattatttgggatactgatcagcaaccattagtttttgaactacct
gaattgagccttcaaaatgctctagtaacaagtgaggtggttgttcacagagatgaaact
atttcattagaatctataattggagctatgactttaactgatgaacttaaagtcaatatt
gtttcattaccgaacactgataaagtagctgatcacctaaccgctaaagttaaggttatt
ttagctgatggctcatttgtcactgtaaatgttccagtcaaagttgtagaaaaagaatta
caaatagctaaaaaggatgctataaaaacaattgatgttctggtaaaacaaaaaatcaaa

```
        gatatagattctaataacgaattaacgtctactcaacgtgaagatgcaaaagctgaaatt
        gaaagattgaaaaagcaagccatcgataaagtgactcattctaaatcgattaaagatatt
        gaaacagtaaaacgaactgattttgaagaaatagatcagtttgatcctaaacgctttacg
        ctaaataaagctaaaaaggatatcattactgatgttaatactcaaatccaaaatgtttc
  5     aaagaaattgaaacaataaaaggtttaacttctaatgaaaaaactcagtttgataaacaa
        ttaactgcactacaaaaagaattttttagaaaaagtcgagcatgctcataatttagtagaa
        ttaaatcaattacaacaagagtttaataatagatatgaacatattttaaaccaagcacat
        ttactaggtgaaaaacatatagcagaacataaattaggatatgttgtagtaaacaaaact
        cagcaaatactaaataatcaatctgcttcttactttataaaacaatgggcacttgataga
  10    attaaacaaattcaactagaaacgatgaattcaattcgtggtgcgcataccgtacaagat
        gtacacaaagcattgttacaaggtatagagcaaatcttgaaagtaaatgtaagtattata
        aatcaatctttcaacgattccttgcataactttaattatcttcattcaaaatttgatgct
        agattaagagaaaaggatgttgcaaaccatatcgtacaaactgaaacattcaaagaagtt
        ctaaaaggaacgggtgttgaaccaggtaaaatcaacaaagaaacacagcaaccaaaactt
  15    cataagaatgataatgatagcctattcaaacatttagttgataatttcggcaaaactgta
        ggtgttattacattaactggtttactttctagtttctggttagttttggctaaaagacgt
        aaaaaagaagaagaaaaacaatcgataaaaaattatcacaaagacattcgtctttcagat
        actgataaaatagatccaattgtaataactaagcgtaaaatagataaagaagaacaaatt
        caaaacgatgacaaacattcaattccagttgctaaacataagaaatctaaagaaaagcaa
  20    ttgagtgaaggagatattcattcaatcccccgtcgttaaacgtaaacaaaacagtgataac
        aaagatacaaaacagaagaaagttacttctaaaaagaagaaaacgcctcaatcaactaaa
        aaagttgtaaaaaccaaaaagcgttctaaaaagtaa
```

Sequence 2586
```
VIKELSQKKRDAINNNTDLTPSQKAHALADIDKTEKDALQHIENSNSIDDINNNKEHAFN
TLAHIIIWDTDQQPLVFELPELSLQNALVTSEVVVHRDETISLESIIGAMTLTDELKVNI
VSLPNTDKVADHLTAKVKVILADGSFVTVNVPVKVVEKELQIAKKDAIKTIDVLVKQKIK
DIDSNNELTSTQREDAKAEIERLKKQAIDKVTHSKSIKDIETVKRTDFEEIDQFDPKRFT
LNKAKKDIITDVNTQIQNGFKEIETIKGLTSNEKTQFDKQLTALQKEFLEKVEHAHNLVE
LNQLQQEFNNRYEHILNQAHLLGEKHIAEHKLGYVVVNKTQQILNNQSASYFIKQWALDR
IKQIQLETMNSIRGAHTVQDVHKALLQGIEQILKVNVSIINQSFNDSLHNFNYLHSKFDA
RLREKDVANHIVQTETFKEVLKGTGVEPGKINKETQQPKLHKNDNDSLFKHLVDNFGKTV
GVITLTGLLSSFWLVLAKRRKKEEEKQSIKNYHKDIRLSDTDKIDPIVITKRKIDKEEQI
QNDDKHSIPVAKHKKSKEKQLSEGDIHSIPVVKRKQNSDNKDTKQKKVTSKKKKTPQSTK
KVVKTKKRSKK*
```

Sequence 2587
Contig_0776_pos_587_147,
is similar to (with p-value 9.0e-29)
>sp:sp|P54170|YPHP_BACSU HYPOTHETICAL 15.9 KD PROTEIN IN ILV
D-THYB INTERGENIC REGION. >gp:gp|L77246|BACYACA_21 Bacillus
subtilis (YAC10-9 clone) DNA region between the serA and kdg
loci. NID: g1256615. >gp:gp|Z99115|BSUB0012_128 Bacillus su
btilis complete genome (section 12 of 21): from 2195541 to 2
409220. NID: g2634478.
```
        atgatgaatggatacgaagcttatatgaaagaacttgcacaacaaatgagagctgaatta
        acagacaatggattcacaagtcttgaaacgagtgatgacgtcaatcagtatatgcaaaat
        atagataatgatgatacaacatttgttgtaatcaactcaacatgcggttgcgctgcagga
        ttagcacgtccagcagctgttgcagttgcagagcaaaatgaagtgaaaccagatcataaa
  50    gtaactgtatttgctggtcaagataaagaagcaacacaaacaatgagagattacatccaa
        caagttccttcaagtccctcatacgcattatttaaaggtcaacatttagttcattttata
        cctcgcgaacatattgaagggcgcgacatcaatgatatagctatggatttaaaagatgct
        tttgatgataattgtcaataa
```

Sequence 2588
```
MMNGYEAYMKELAQQMRAELTDNGFTSLETSDDVNQYMQNIDNDDTTFVVINSTCGCAAG
LARPAAVAVAEQNEVKPDHKVTVFAGQDKEATQTMRDYIQQVPSSPSYALFKGQHLVHFI
PREHIEGRDINDIAMDLKDAFDDNCQ*
```

Sequence 2589
Contig_0777_pos_367_684,
putative peptide of unknown function
atgaggagaatatatatgaaaagattattaggtacattaattgctgctacactagtgtta
agtgcttgtagccaaaacgacactaaggaagatgaaaataaaaagtcagaaaatactact
gaaaagaaatctgacgataaaaaagacaaaaaaactaatgaggataaaaagtctggagaa
caaaagaaatctcaagaaaaaaagaataacaagtcaatgcaagaatctgctacaaatgaa
caggttcaatctcaacaacaaacttcacaaggtgtaccatcaaaatgctctttgccgatg
actgttttatcaatttga

Sequence 2590
MRRIYMKRLLGTLIAATLVLSACSQNDTKEDENKKSENTTEKKSDDKKDKKTNEDKKSGE
QKKSQEKKNNKSMQESATNEQVQSQQQTSQGVPSKCSLPMTVLSI*

Sequence 2591
Contig_0777_pos_6354_6046,
putative peptide of unknown function
atgactgcattttaattcttaatttattgactatgaaagaggcttcaatatcgtctatt
attgtacgaactgtcattgcagctattgttttctttgtcatttatatcattgtatttaca
attttaagttcgtcagaacgtaaacttatttatggtacaactttgcctattgcgctttt
atatgccttatattcggagcaattttttttcactccgcgtataggtatcattgccggacta
attataggtgtgtttgctggtgtcatatgggagttcttaaatagaaaaaatggaggtcgc
tcatcttga

Sequence 2592
MTAFLILNLLTMKEASISSIIVRTVIAAIVFFVIYIIVFTILSSSERKLIYGTTLPIALF
ICLIFGAIFFTPRIGIIAGLIIGVFAGVIWEFLNRKNGGRSS*

Sequence 2593
Contig_0777_pos_4678_1709,
is similar to (with p-value 0.0e+00)
>sp:sp|P40815|T3RE_SALTY TYPE III RESTRICTION-MODIFICATION S
YSTEM STYLTI ENZYME RES (EC 3.1.21.5). >pir:pir|JN0658|JN065
8 restriction endonuclease (EC 3.1.-.-) - Salmonella typhimu
rium
atgaaaatattcttagaagaactaattcaccaacaacaagcagttaaaaaaataatagat
actttcacaggaatcgaaaagtatttaacttcaaaaaattgtgacaatgagttcgctaat
aatttaataataaatagatattcagaagaagctaacatagatataaaaatggagaccggg
acaggtaaaacgtatgtttatactaaaatgatgtatgaattacataaaaaatttgggatt
tttaaatttatattagtagttccaagtcctgcaattaaagagggagcaaagaattttta
actagtttatcaactaaaagacatttccaagaaacatacggaaatgttgaaatagaaata
aatacaataaataaaggcgatttaatactagatcaggcagaaaaattttccacctcat
ttgagtaactttattgaaagcagtaatttaaatgctaatcaaattcaagtattacttatt
aatgcaggtatgttaaattcatcaaatatgacaaaggttgattacgaccaaactttatta
agtaattactcaaaccctatagaagctttaaaagctaccaaatctgtagttataatagac
gaaccccataggtttcctagggataagaagaactacaaatctatagaaaatcttgaacct
caaatgattgttaggtttggtgctactttcctgaagtaaaaaaaggaacaggaaaaaaa
gctgtatatattaaagattattatcgaggtaggcctcaatttgaattgaatgctgttgac
agctttaatcagggacttgtaaagggaatagatatttactatccaaatttgactccagaa
caagcgaaaaatagatataccattgatagtgtaaaagccaaagagatagttttgaaaaaa
ggaaaaaatalaatggacgctaggtataggcgagaatttagcaaatatagattcgttattt
gaaggtgacttaagttattctggtgctaaaactttatctaatgacttagaaattagtaaa
ggtatggatttattgcctggcacatttacaactaattaccaggaattaattattaatgat
gctatcaatcaacattttgagcatgaaatcaataattttatgagagataacctaaaagaa
aatttttcaaccaaaagtaaaaacactatcattattttttattgattctattaggtcatat
agaaacaaagagggctggttgaaacagacttttgagagactcttaaaagttaaacttaga
aaattaataaaagaatttgaagttaaaaaattaccacgagagattgaatacttagatttt
ttaagagtgacatatgatagtcttaatagtgagaatcaaatggtccatgcagggtatttt
ggtgaagacagaggttccggtgatgaagcaatacaggctgaagttaatgatatacttaaa

aataaagaaaaaatgctgagttttaaagacgagaatggtaactggattacaagaagattt
ttattttcaaaatggacattaagagaaggttgggataaccctaatgttttcgttattgcg
aaattgagaacttctggttctgaaaatagcaagatacaggaagtaggaagaggtttgcgt
ttacctgttgatgagaatggtcacagattaacacaagatgaatttccaagtagactttca
5 tttttaattggatatgatgaaaaggatttcgcagaaaaattaattggagaaataaattct
gatgtagacgttaaattaagtgaagataaattaacagatgaaatgattaataaaatagtt
gaacatagaaagcaagtggatcctcattataatgatgaagtgttattagaacaattagat
gaacgaaatctaattaataggaaaaatgagtttaaaacagatgttgaaattgatggagtt
aaaaaatcaggatttgaatggctgttagaattatatccagaagttaatgcatctaaatta
10 aatgcagataaagtcagagatatgaaaaaaaatccacccaatttaaaagtgaagttaaat
aaagagaattggaataaattgagattcttatgggaaaatttatctaaaagatatatgtta
gaattcaaaaagatgagtgaagatgatctgtatttgtttgttgaaagattgttgaacgac
gatgatttatttgttaaacaacaaccagaaagaatacatcaatctttggaaaaagatgat
gaaggtaaacaagttatcaaggaatctatttcagaatataattatagaaatgaatttatt
15 tatatgaactatggtaagttttaaaacaaattgtacttaaaactaatgtaccaatatct
ataatgcataaaaatttactttccgttttaaaagataaatataatagtgatgaacgattc
ttgagtgaattaagtttaaataatataataagagaatttaacaagcgttttgaagaaaaa
tatagtcaaagttatgaatataagaaattagattttctgctactacaaccatttatgat
tcagaaatatcagagtttaaagattgggtagatgcaaattatttaggtactaacgttgaa
20 aataacattcaaactcaaaaaagatttttatatgaaagaccaccagttagatatgatagt
gtaacacctgagttagagttgttaaaaagaaattacgataaaaatgtaactgtatttggt
aatttgcctaaaaaagcgatacaagttcctaaatatactggtggcactactacgcctgat
tttgtctatatgatagaaactgatgaacaagatgcaaaataccttattgttgaaacaaaa
gcagaaaacatgagactaggagataaaagtattggtgaaatacaaaaaaaattctttaac
25 acattagataatttgaatattaaatatcaattagctactagcgcgcaagatgtttataat
gaaattaaaaaattagatgattcaaagtga

Sequence 2594
MKIFLEELIHQQQAVKKIIDTFTGIEKYLTSKNCDNEFANNLIINRYSEEANIDIKMETG
30 TGKTYVYTKMMYELHKKFGIFKFILVVPSPAIKEGAKNFLTSLSTKRHFQETYGNVEIEI
NTINKGDFNTRSGRKIFPPHLSNFIESSNLNANQIQVLLINAGMLNSSNMTKVDYDQTLL
SNYSNPIEALKATKSVVIIDEPHRFPRDKKNYKSIENLEPQMIVRFGATFPEVKKGTGKK
AVYIKDYYRGRPQFELNAVDSFNQGLVKGIDIYYPNLTPEQAKNRYTIDSVKAKEIVLKK
GKNKWTLGIGENLANIDSLFEGDLSYSGAKTLSNDLEISKGMDLLPGTFTTNYQELIIND
35 AINQHFEHEINNFMRDNLKENFQPKVKTLSLFFIDSIRSYRNKEGWLKQTFERLLKVKLR
KLIKEFEVKKLPREIEYLDFLRVTYDSLNSENQMVHAGYFGEDRGSGDEAIQAEVNDILK
NKEKMLSFKDENGNWITRRFLFSKWTLREGWDNPNVFVIAKLRTSGSENSKIQEVGRGLR
LPVDENGHRLTQDEFPSRLSFLIGYDEKDFAEKLIGEINSDVDVKLSEDKLTDEMINKIV
EHRKQVDPHYNDEVLLEQLDERNLINRKNEFKTDVEIDGVKKSGFEWLLELYPEVNASKL
40 NADKVRDMKKNPPNLKVKLNKENWNKLRFLWENLSKRYMLEFKKMSEDDLYLFVERLLND
DDLFVKQQPERIHQSLEKDDEGKQVIKESISEYNYRNEFIYMNYGKFLKQIVLKTNVPIS
IMHKNLLSVLKDKYNSDERFLSELSLNNIIREFNKRFEEKYSQSYEYKKLDFSATTTIYD
SEISEFKDWVDANYLGTNVENNIQTEKRFLYERPPVRYDSVTPELELLKRNYDKNVTVFG
NLPKKAIQVPKYTGGTTTPDFVYMIETDEQDAKYLIVETKAENMRLGDKSIGEIQKKFFN
45 TLDNLNIKYQLATSAQDVYNEIKKLDDSK*

Sequence 2595
Contig_0777_pos_1393_623,
is similar to (with p-value 0.0e+00)
50 >gp:gp|U92974|LLU92974_13 Lactococcus lactis unknown gene, p
artial cds, and HisC (hisC), unknown, HisG (hisG), unknown,
HisB (hisB), unknown, HisH (hish), HisA (hisA), HisF (hisF),
HisIE (hisIE), unknown, unknown, LeuA (leuA), LeuB (leuB),
LeuC (leuC), LeuD (leuD), unknown, IlvD (ilvD), IlvB (ilvB),
55 IlvN, IlvC (ilvC), IlvA (ilvA), AldB (aldB) and aldR (aldR)
genes, complete cds. NID: g2565137.
atggattatagagtactactttattataaatatgtaactatagatgaccctgaaactttt
gcagccgaacatttgaaattttgtaaggaacatcatttaaaaggaagaatactagtttca
acggaaggcattaatggaacattatctggaacaaaagaagatactgataaatatatagag

catatgcatgcagatagtcgttttgctgatttaacttttaaaattgatgaagctgaaagt
catgcgtttaaaaagatgcacgtgcgtccaagacgtgaaattgttgcacttgacttagaa
gaagatattaatccacgtgaaattaccggtaaatactattctcctaaagaatttaaagcc
gcactagaagatgaaaatactgttatattagatgctcgaaatgattatgaatacgattta
5   ggacatttccgtggagctattcgtcctgatataacacgattccgtgacttacctgaatgg
gtgcgtaataataaagaacaactcgacggaaaaaatattgtcacatattgtacaggtggc
attcgttgtgaaaaattttctggttggttagtaaaagaaggatttgaaaacgtaggtcag
ttgcatggtggtattgctacatacggtaaagaccctgaaactaaagggctatattgggat
ggtaagatgtatgtatttgatgaacgtattagtgtcgatgtgaatcaaattgataaaaca
10   gtcatcggcaaagagcattttgatggtacaccttgtgaagtttgttgttga

Sequence 2596
MDYRVLLYYKYVTIDDPETFAAEHLKFCKEHHLKGRILVSTEGINGTLSGTKEDTDKYIE
HMHADSRFADLTFKIDEAESHAFKKMHVRPRREIVALDLEEDINPREITGKYYSPKEFKA
15   ALEDENTVILDARNDYEYDLGHFRGAIRPDITRFRDLPEWVRNNKEQLDGKNIVTYCTGG
IRCEKFSGWLVKEGFENVGQLHGGIATYGKDPETKGLYWDGKMYVFDERISVDVNQIDKT
VIGKEHFDGTPCEVCC*

Sequence 2597
20   Contig_0778_pos_324_815,
is similar to (with p-value 5.0e-52)
>sp:sp|P38424|YSXC_BACSU HYPOTHETICAL GTP-BINDING PROTEIN IN
LONA-HEMA INTERGENIC REGION (ORFX). >pir:pir|S45102|S45102
hypothetical protein X - Bacillus subtilis >gp:gp|X76424|BSL
25   ONLA_3 B.subtilis lon gene for protease La. NID: g496556. >g
p:gp|Z99118|BSUB0015_84 Bacillus subtilis complete genome (s
ection 15 of 21): from 2795131 to 3013540. NID: g2635200. >g
p:gp|Z75208|BSZ75208_85 B.subtilis genomic sequence 89009bp.
NID: g1769994.
30   atgaatataaattttaataatattaacttaattataagtgctgtaaaaaaagcacagtat
cctgacactggattaacagaagtagcgttaagtggacgctcaaatgtagggaaatctaca
tttattaatagtatgattgggcgtaaaaatatggcgagaacgtcacaacaacctggtaag
acacagacattgaatttctataatatagatgaacaacttattttttgttgatgtaccagga
tatggatacgctaaagtaagtaaagttcaacgagaaaatttggtaaaatgattgaagaa
35   tatattacacaacgagagaatttaaaacttgttattcaacttgtcgatttaagacatcaa
cctactgaagatgatgtgcttatgtacaattatcttaaacattttgatataccaacactt
gtaatatgtactaaagaagataaaattgccaaaggaaaagtacgagggtccagagaggct
ccgaagatatga

40   Sequence 2598
MNINFNNINLIISAVKKAQYPDTGLTEVALSGRSNVGKSTFINSMIGRKNMARTSQQPGK
TQTLNFYNIDEQLIFVDVPGYGYAKVSKVQREKFGKMIEEYITQRENLKLVIQLVDLRHQ
PTEDDVLMYNYLKHFDIPTLVICTKEDKIAKGKVRGSREAPKI*

45   Sequence 2599
Contig_0778_pos_2037_3047,
is similar to (with p-value 8.0e-80)
>sp:sp|P41006|PYRP_BACCL URACIL PERMEASE (URACIL TRANSPORTER
). >pir:pir|S38893|S38893 uracil transport protein - Bacillu
50   s caldolyticus >gp:gp|X76083|BCPYRQP_2 B.caldolyticus (DSM40
5) pyrR, pyrP and pyrB (partial) genes. NID: g431229.
atgctggttgcattatttatgagtggattaatgtacgtgattataggtattttcattaaa
ttgagtggaacacattggttaatgcacttgttaccaccagtagttgtcggaccagtaata
atggtcattgggttaagtttagctcctacagcagtaaacatggccatgttcgaaaattct
55   gctgaaatgaaagggtataacttaagttacttaattgttgctttgattacattagcagta
accatcatcgtccaaggattcttcaaaggatttttatcactaatacctgtacttataggt
attatagtgggatatattgtatccattttcatgggcatagttaaatttgctccaatagca
caagcgaaatggatagattttcctcatatttatctaccatttaaagattacacaccatct
tttcatttaggactcattctcgtgatgatacccgtggtgtttgtgacggtaagtgaacat

attggtcatcaaatggtaattaataaaatagtaggacgcaatttctttgaaaatccaggt
ttagataaatcaatcattggtgatggtgtttcaactatgtttgcaagtatgataggaggt
cctcctagtacaacttatggtgaaaatataggtgtactagcgatcaccaaaatatatagt
atttacgttattggtggtgcggcagttatagctatcattcttgcatttattggtaagttc
5   actgctttaatatcttcaatacctacgccagtgatgggtggtgtctcaattttattattc
ggtattatagcagctagtggtttaagaatgcttgttgaaagtcaagtagatttcgcaagc
aatcgcaacttggttatagcatcagttgtgcttgttgtcgggattggtaatcttcttatc
aatttaaaaggcataggtatcaatttacaaattgaaggaatggcattatcagcacttca
ggaataatattaaatttaattttgccaaaagataaaaaccaaataaattaa

10

Sequence 2600
MLVALFMSGLMYVIIGIFIKLSGTHWLMHLLPPVVVGPVIMVIGLSLAPTAVNMAMFENS
AEMKGYNLSYLIVALITLAVTIIVQGFFKGFLSLIPVLIGIIVGYIVSIFMGIVKFAPIA
QAKWIDFPHIYLPFKDYTPSFHLGLILVMIPVVFVTVSEHIGHQMVINKIVGRNFFENPG
15 LDKSIIGDGVSTMFASMIGGPPSTTYGENIGVLAITKIYSIYVIGGAAVIAIILAFIGKF
TALISSIPTPVMGGVSILLFGIIAASGLRMLVESQVDFASNRNLVIASVVLVVGIGNLLI
NLKGIGINLQIEGMALSALSGIILNLILPKDKNQIN*

Sequence 2601
20 Contig_0778_pos_3073_3954,
is similar to (with p-value 2.0e-90)
>sp:sp|P05654|PYRB_BACSU ASPARTATE CARBAMOYLTRANSFERASE (EC
2.1.3.2) (ASPARTATE TRANSCARBAMYLASE) (ATCASE). >pir:pir|A25
015|OWBSAC aspartate carbamoyltransferase (EC 2.1.3.2) catal
25 ytic chain - Bacillus subtilis >gp:gp|M13128|BACPYRB_1 B.sub
tilis pyrB gene encoding aspartate transcarbamoylase, comple
te cds. NID: g143383. >gp:gp|M59757|BACPYROP_3 Bacillus subt
ilis pyrimidine biosynthetic (pyr) gene cluster (pyrR, pyrP,
pyrB, pyrC, pyrAA, pyrAB, pyrD, pyrF and pyrE) genes, compl
30 ete cds. NID: g387576. >gp:gp|Z99112|BSUB0009_20 Bacillus su
btilis complete genome (section 9 of 21): from 1598421 to 18
07200. NID: g2633902.
atgaacacttattatcaatggagcatttatctaattcagaaatttatgatttaattact
atcgcttgccaattcaaatctggtgagcgaccattacctcaatttaacggtcaatacgta
35 tcaaacttattcttcgaaaattcaacgcgaacaaagtgtagctttgagatggcagaacaa
aaattaggattaaaacttattaattttgaaacaagtacatcatctgtaaaaaagggtgag
tcactttatgacacatgtaaaacacttgaaagtataggtgttgatttacttgtcatacgt
cactcccaaaattcttattacgaagaactggatcaattaaatattccaattgctaatgca
ggtgatggaagtggacaacatcctactcagagtttattagacataatgacaatatgaa
40 gaatatggttcgtttgaaggtttgaatattctaatatgtggggacattaaaaattctcgt
gtcgcaagaagtaattatcatagtttaacatcattaggtgccaacgtaatgttctcaagt
ccaaaagaatgggtagataatacattagaggcgccttatgttgaaattgatgaagtcatt
gataaagtagatattgttatgttgcttagagttcaacatgaaagacatggaatttcaggt
gaagctaactttgctgctgaagaatatcatcaacaatttggtttaacacaggctagatat
45 gataaattaaaagaggaagccattgtaatgcatccagctcctgaaatagaggtgttgaa
attaaaagcgagctagttgaagcacctaagtctcgaatatttaagcagatggaaaatgga
atgtatttaagaatggcagtaataagtgcgcttttacaatag

Sequence 2602
50 MEHLLSMEHLSNSEIYDLITIACQFKSGERPLPQFNGQYVSNLFFENSTRTKCSFEMAEQ
KLGLKLINFETSTSSVKKGESLYDTCKTLESIGVDLLVIRHSQNSYYEELDQLNIPIANA
GDGSGQHPTQSLLDIMTIYEEYGSFEGLNILICGDIKNSRVARSNYHSLTSLGANVMFSS
PKEWVDNTLEAPYVEIDEVIDKVDIVMLLRVQHERHGISGEANFAAEEYHQQFGLTQARY
DKLKEEAIVMHPAPVNRGVEIKSELVEAPKSRIFKQMENGMYLRMAVISALLQ*

55

Sequence 2603
Contig_0778_pos_3972_5249,
is similar to (with p-value 0.0e+00)
>sp:sp|P46538|PYRC_BACCL DIHYDROOROTASE (EC 3.5.2.3) (DHOASE

638

). >pir:pir|S34319|S34319 dihydroorotase (EC 3.5.2.3) - Baci
llus caldolyticus >gp:gp|X73308|BCPYR_2 B.caldolyticus pyrim
idine biosynthesis genes. NID: g312439.
atgaaattaattaaaaacggaaaaatcttaaaaaacggtatcctaaaagacacagaaatt
5   ttaatcgacggtaaacgtattaaacaaattagtagtaaaattaatgcttcatcttcaaat
attgaagttattgatgcaaaaggaaatttaattgctcccggtttgtagatgttcatgtg
cacctacgtgaaccaggtggtgaacataaagaaacaattgaaagtggtacaaaagccgct
gcaagaggtggtttttactacagtatgtcctatgcctaatacaagacctgtaccagataca
gttgaacatgttagagaattaagacaacgaatttctgaaacagcacaagttagggtgttg
10   ccttatgctgctattactaagagacaagcaggtactgaacttgttgattttgaaaaatta
gcactagaaggtgtgtttgcatttactgacgatggtgtgggagttcaaacagcaagtatg
atgtatgctgctatgaagcaagctgcaaaagttaaaaaaccgattgtcgcacactgtgaa
gataatagcttaatctatggtggtgcaatgcataaaggtaaacgtagtgaagaattaggc
atacctggtattccaaatattgctgaatctgtacaaattgctagagatgtattattggct
15   gaagcaactggttgtcactatcatgtgtgtcatgtttcaactaaggaaagtgttcgagta
atcagagacgctaaaaaagctggtatccatgtaacagcagaagttacaccacatcattta
ttattaactgaaaatgatgttcctggcgatgattcaaactacaaaatgaatccaccatta
agaagtaatgaagatagagaagcactttttagaaggcttattagatggaacaattgattgt
attgcaacggatcatgcacctcacgctaaagaagaaaaagcacaacctatgacaaaagca
20   cctttcgqcatcgtaggtagtgaaacagcattcccattactttatacacactttgt.aaga
cgaggtaattggtcactgcaacaattagttgattatttcactattaaaccagctactatt
ttcaacttaaattatggaaaattacacaaagatagttacgctgatttaacaataattgat
cttaatactgaaaaagaaatcaaaagtgaagatttcttatctaaagctgataacactcca
tttattggtgaaaaagtttatggaaatccaacactaacaatgcttaaaggtgaagtagta
25   ttcgaggaggaaaagtag

Sequence 2604
MKLIKNGKILKNGILKDTEILIDGKRIKQISSKINASSSNIEVIDAKGNLIAPGFVDVHV
HLREPGGEHKETIESGTKAAARGGFTTVCPMPNTRPVPDTVEHVRELRQRISETAQVRVL
30   PYAAITKRQAGTELVDFEKLALEGVFAFTDDGVGVQTASMMYAAMKQAAKVKKPIVAHCE
DNSLIYGGAMHKGKRSEELGIPGIPNIAESVQIARDVLLAEATGCHYHVCHVSTKESVRV
IRDAKKAGIHVTAEVTPHHLLLTENDVPGDDSNYKMNPPLRSNEDREALLEGLLDGTIDC
IATDHAPHAKEEKAQPMTKAPFGIVGSETAFPLLYTHFVRRGNWSLQQLVDYFTIKPATI
FNLNYGKLHKDSYADLTIIDLNTEKEIKSEDFLSKADNTPFIGEKVYGNPTLTMLKGEVV
35   FEEEK*

Sequence 2605
Contig_0778_pos_5250_0,
is similar to (with p-value 5.0e-51)
40   >sp.:sp|P25993|CARA_BACSU CARBAMOYL-PHOSPHATE SYNTHASE, PYRIM
IDINE-SPECIFIC, SMALL CHAIN (EC 6.3.5.5) (CARBAMOYL-PHOSPHAT
E SYNTHETASE GLUTAMINE CHAIN). >pir:pir|E39845|E39845 carbam
oyl-phosphate synthase (glutamine-hydrolyzing) (EC 6.3.5.5),
pyrimidine-repressible, small chain - Bacillus subtilis >gp
45   :gp|M59757|BACPYROP_5 Bacillus subtilis pyrimidine biosynthe
tic (pyr) gene cluster (pyrR, pyrP,pyrB, pyrC, pyrAA, pyrAB,
pyrD, pyrF and pyrE) genes, complete cds. NID: g387576. >g
p:gp|Z99112|BSUB0009_22 Bacillus subtilis complete genome (s
ection 9 of 21): from 1598421 to 1807200. NID: g2633902.
50   atgcttgaaaaacgttatcttgtactggaagatggctcttattacgaaggatatcgctta
gggtcagatgacttatctataggcgaaattgtattcaacactgctatgacggggtaccaa
gaaacaatctctgacccgtcttacacaggtcaaatcataactttacgtacccactaatt
ggaaactatggtattaatcgcgatgattttgaatcattaacacctaaattaaatggggta
gtagtaaaggaagcaagtacacaccctagtaactttagacaccaaaaaactttacacgaa
55   acacttgctcaatatcatattcctggtatatcggggtagatactagaagtattactcgt
aaaattagaaattatggtgtttaagagctggatttacagataataaagataacattcag
gaacttgttgaacagttgaaaactgctgaattacctagagatgaagttcaaacggtttct
acaaaaacaccatatgtatcaacaggttccgatttaagcgtcgttttactcgactttggt
aaaaagc

Sequence 2606
MLEKRYLVLEDGSYYEGYRLGSDDLSIGEIVFNTAMTGYQETISDPSYTGQIITFTYPLI
GNYGINRDDFESLTPKLNGVVVKEASTHPSNFRHQKTLHETLAQYHIPGISGVDTRSITR
KIRNYGVLRAGFTDNKDNIQELVEQLKTAELPRDEVQTVSTKTPYVSTGSDLSVVLLDFG
KKX

Sequence 2607
Contig_0779_pos_3369_1672,
is similar to (with p-value 0.0e+00)
>sp:sp|P18255|SYT1_BACSU THREONYL-TRNA SYNTHETASE 1 (EC 6.1.
1.3) (THREONINE--TRNA LIGASE) (THRRS). >pir:pir|B37770|YSBST
1 threonine--tRNA ligase (EC 6.1.1.3) 1 - Bacillus subtilis
>gp:gp|AF008220|AF008220_195 Bacillus subtilis rrnB-dnaB gen
omic region. NID: g2293135. >gp:gp|M36594|BACTRNASB_1 B.subt
ilis threonyl-tRNA synthetase (thrSv) gene, complete cds. NI
D: g143765. >gp:gp|Z99118|BSUB0015_160 Bacillus subtilis com
plete genome (section 15 of 21): from 2795131 to 3013540. NI
D: g2635200. >gp:gp|Z75208|BSZ75208_5 B.subtilis genomic seq
uence 89009bp. NID: g1769994.
atggcacaagcattaaaacgttatacggagacgttaaatttggagttggacctgtaata
gaaggcggattctattatgattttgatatggatgataaggtttcatcggatgattttgat
aaaattgagaaacaatgaaacaaattgtgaacgaaaatcataaaattgtaagagaagta
gttagtaaagaaaaagcaaaagacttcttcaaggatgacccttataaattagaacttatt
gatgcaattcctgaagatgagagtgtaacactttatactcaaggtgaatttactgattta
tgtcgaggtgtacacgtaccttctacttctaaaattaaagagttcaaactattatctaca
gctggtgcttattggcgtggaaatagtgataataaaatgttacaacgaatttatgtgtaca
gcattctttgacaaaaaagatttgaaagcacatctaaaaatgttggaagaacgtcgtgag
cgtgatcatcgtaaaattggtaaagatttagaattgtttacaaacaatcaactcgttggt
gctggtttaccattatggttaccaaatggtgctacaatacgtagggaaatagaacgttat
attgtcgataaagaagtaagtatgggatacgatcatgtttacacaccagtattagccaat
gttgatttatataaaacatctggtcactgggatcattatcaagaagatatgttccgagca
atgaagttagatgaagacgaagcaatggtcttaagaccaatgaactgtccacatcatatg
atgatttataaaaacaaacctcattcttatcgcgaattacctatacgtattgctgaattg
ggtactatgcatcgttacgaagcaagtggtgcagtatcaggtttacaacgtgttcgagga
atgacattgaatgattcccatattttcgttagacctgatcaaattaaagaagaatttaaa
cgtgtagttaatatgattcaagatgtgtacaaagattttggttttgaagattatcgcttc
agattgagttatagagatcctgaagataagcataagtactttgatgatgatgaaatgtgg
gaaaaagctgaatccatgcttaaagaagcatcagatgaattaggtttaacttatgaagaa
gctattggtgaggcagcattctatggacctaagttagatgttcaagtaaaaacagctatg
ggaaaagaagaaactctatcaacagcacaacttgattttcttttaccagaacgttttgac
ttaacgtacattggtcaagatggagaacaacatcgtcctgtagttatacaccgtggtgta
gtttctactatggaacgtttttgttgcatttttaacagaagaaacaaaaggtgcatttcca
acttggttggcgcctatgcaagttgaaattattcctgtaaatatagatttacattatgat
tatgcaagacttttacaagatgaactaaaatcgcaaggtgtccgcgttgaaattgatgac
cgtaatgaaaaaatgggatataaaattcgtgaagctcaaatgaaaaaaataccttatcag
attgttgtaggtgaccaagaagtagagaatcaagaagtaaatgtaagaaaatatggttct
gaaaaacaagaatcagttgaaaaagatgaatttatttggaatgttattgatgaaatccgt
ttgaaaaagcatagataa

Sequence 2608
MAQALKRLYGDVKFGVGPVIEGGFYYDFDMDDKVSSDDFDKIEKTMKQIVNENHKIVREV
VSKEKAKDFFKDDPYKLELIDAIPEDESVTLYTQGEFTDLCRGVHVPSTSKIKEFKLLST
AGAYWRGNSDNKMLQRIYGTAFFDKKDLKAHLKMLEERRERDHRKIGKDLELFTNNQLVG
AGLPLWLPNGATIRREIERYIVDKEVSMGYDHVYTPVLANVDLYKTSGHWDHYQEDMFPA
MKLDEDEAMVLRPMNCPHHMMIYKNKPHSYRELPIRIAELGTMHRYEASGAVSGLQRVRG
MTLNDSHIFVRPDQIKEEFKRVVNMIQDVYKDFGFEDYRFRLSYRDPEDKHKYFDDDEMW
EKAESMLKEASDELGLTYEEAIGEAAFYGPKLDVQVKTAMGKEETLSTAQLDFLLPERFD
LTYIGQDGEQHRPVVIHRGVVSTMERFVAFLTEETKGAFPTWLAPMQVEIIPVNIDLHYD

640

YARLLQDELKSQGVRVEIDDRNEKMGYKIREAQMKKIPYQIVVGDQEVENQEVNVRKYGS
EKQESVEKDEFIWNVIDEIRLKKHR*

Sequence 2609
Contig_0781_pos_5154_4732,
is similar to (with p-value 6.0e-49)
>gp:gp|D88209|D88209_1 Bacillus licheniformis DNA for Pz-pep
tidase, complete cds. NID: g1651215.
atgtttgctgaatttgaacataaaatacatcaaatagaagaagctggggagccgttaacg
ccaaatcgtatgaatgaagaatatgctaaactgaacaaactatattttggtgaagcagta
gaaactgacgatgatattagtaaagaatggtcacgtattcctcatttctatatgaattat
tatgtatatcaatacgcaactggttatagtgcagctcaaagtttaagtcatcaaatttta
actgagggtcaacctgctgttgaacgatatatcaatgaattcttaaaaaagggtagctca
aactatccgattgaaattttaaaaaatgcaggtgttgacatgacaacacctcaaccaata
gaggaagcttgtgaagtattcgaacaaaaattagatgcttttgaaaagttaatgaaagct
tag

Sequence 2610
MFAEFEHKIHQIEEAGEPLTPNRMNEEYAKLNKLYFGEAVETDDDISKEWSRIPHFYMNY
YVYQYATGYSAAQSLSHQILTEGQPAVERYINEFLKKGSSNYPIEILKNAGVDMTTPQPI
EEACEVFEQKLDAFEKLMKA*

Sequence 2611
Contig_0781_pos_3771_3247,
putative peptide of unknown function
atgtatgttggtacgagtaaaagtagtgtttaaatctaactgcagttcttaaaggtaaa
gatatttatcatgtatatgctgaatataagtctccatattataagcaatatggtaaatca
gaagcccctacaatatatgatgataatattacaagtcaatcagagttaaagaagaaatta
aaagaaacacttgatgacataccaacaatcgaagtagcaacgaattatttaggattagaa
agtattcatgaaaataatactattcgatttatacacaaacctatcggatttaatactgat
ttaaaagttgtcaaacttactgaatatcaccccccttgtttcgcagcctattgaagtggaa
ttcagtaatgcacagaaagatattataaaaatgcaatcacagttcaatcgtaggttaaga
aaggttaataatcttatgaaaaaaggattcaaaactagtgactattctttaaatgtgtta
gaggaatataacgaaacagtaggaagtgtattgattgatgagtaa

Sequence 2612
MYVGTSKSSVLNLTAVLKGKDIYHVYAEYKSPYYKQYGKSEAPTIYDDNITSQSELKKKL
KETLDDIPTIEVATNYLGLESIHENNTIRFIHKPIGFNTDLKVVKLTEYHPLVSQPIEVE
FSNAQKDIIKMQSQFNRRLRKVNNLMKKGFKTSDYSLNVLEEYNETVGSVLIDE*

Sequence 2613
Contig_0781_pos_3065_1200,
putative peptide of unknown function
atgttattaactttagactttcctattcaaataggacacacatttagaaccaagatgata
aataattttagaacaatacttaattattataatgaattagatcatcagcatcgcgcacac
acagaaactaagcatcatgcacatcaagccatgcaggttgattatagaaatacaaacgtt
tctgcatttttagattatcttaacggtaatattaatgggcttgtttttaggagcaaatgga
gacggtatagctgaaacaaaacaagccagagtatcaatagatggtaccgtacatcccttg
ttgcaagaaaggctgcttcatgactttttaggaattaacagaaaattagataaagaaata
cattctaatggtgcagttgatttttatttggaatcctccatatataccaggaaatagattg
ggagaaaatgggacaccaaataattgggaaccagaagcccatattgaagcgttttttaaac
cctttagttgataatcaatacgttacaaaagaagttataggagaagatacatcaggaaaa
tataatgtgtacaaatttacgtttgaaccacaaaattacaataaaacgttacttattact
tcatgtatacacggtaatgaaactactggattttttgatatgtgccatatactcaatcta
ttggtcaatcaatgggaaaagtatcctcaattaacttacttaagaaaaaatgtacgttta
atttatgttcctatggttaacccgtggggattcgcaaatcaagaaagagagaatgtgaac
aatgtagatttaaacagaaattttgattataactggaaggcaggtaaagggacagatcct
gataaatctaacttcaaaggtaaaagtcctttttctgaaaaagaatcacaaaatatgcgt
agcttagttcaaagtatagataatttaactgctcacttagatttgcatgatattatttca

```
       gtaaataatgattactgtttattttatccgcgttgggccaatcaaaaaaataataatatg
       actcatcttattaacaatttaaaaagtaacggagacctcgttgtttggggttccagtaca
       ttatcatcttttagtaattgggtaggaatccgaaataaaacaacgtcatatctttcagaa
       ataaatgaaaaacgtgtcggtgaaaagaaaagtcccgaagaaatgagacgttcagtacgc
5      tgggtaggtaatgtaattttttagaatggcacaatttgaatcttatcaaaatggtcaaaca
       tcattagatcctttcattaaagtgatggtatatgatgatagatttaacaataaaacatct
       gaagtcattaccctacgtgcagaaaggaatgaatggcaacgtataatgatgagtcagcag
       cgtttcaaagtttttagcaaatggatttgtagagctctatggatatgtgactataaacgtt
       gatagagatgtcacagtggggattaatcctaatattgttcagaattatcatccattcttt
10     ggatttaataaaagtagaaaacgtaatttattttcaattgaacatagactcaacaaagga
       aatacaactttccctatttacgctgctgctggagttcaaatgtcgacgattactgaacca
       ggtacaaaacgtactgatacagtaatgccggtactagatgttaagaaaaaaggtgctggt
       attgtaacaatcaaacaaattaaattatttgcgaagttcactcctacgcattctgctaat
       tccattcagatattaaaatctggagaatattattttttatttttttctaacgctttgcgat
15     atcgatcatcattatcccacatttcttgaatttgtggttctttattcggtaagcccctc
       tcatag


       Sequence 2614
       MLLTLDFPIQIGHTFRTKMINNFRTILNYYNELDHQHRAHTETKHHAHQAMQVDYRNTNV
20     SAFLDYLNGNINGLVLGANGDGIAETKQARVSIDGTVHPLLQERLLHDFLGINRKLDKEI
       HSNGAVDFIWNPPYIPGNRLGENGTPNNWEPEAHIEAFLNPLVDNQYVTKEVIGEDTSGK
       YNVYKFTFEPQNYNKTLLITSCIHGNETTGFFDMCHILNLLVNQWEKYPQLTYLRKNVRL
       IYVPMVNPWGFANQERENVNNVDLNRNFDYNWKAGKGTDPDKSNFKGKSPFSEKESQNMR
       SLVQSIDNLTAHLDLHDIISVNNDYCLFYPRWANQKNNNMTHLINNLKSNGDLVVWGSST
25     LSSFSNWVGIRNKTTSYLSEINEKRVGEKKSPEEMRRSVRWVGNVIFRMAQFESYQNGQT
       SLDPFIKVMVYDDRFNNKTSEVITLRAERNEWQRIMMSQQRFKVLANGFVELYGYVTINV
       DRDVTVGINPNIVQNYHPFFGFNKSRKRNLFSIEHRLNKGNTTFPIYAAAGVQMSTITEP
       GTKRTDTVMPVLDVKKKGAGIVTIKQIKLFAKFTPTHSANSIQILKSGEYYFYFFLTLCD
       IAHHYPTFLEFVVLYSVSHLS*
30

       Sequence 2615
       Contig_0784_pos_1894_2202,
       putative peptide of unknown function
       gtcactaaagcttctaaagattttgtagtaaacattaacatatgtcctccgaatagatct
35     tggatagcatttgtaatcagtactccaggtacaatcggcatgaccgctgcaatgataatt
       gtagctaaatctccgctaggaacaaatgcatgtccaattacagaaataatacctattacc
       aaagaacctatgaattctggaataaaattgtgcgtgtagctttcgatctaatatttctact
       actaagtatccaatcgttccagctaatacagctgtgatgatatcaaccagacgacctccc
       tgtagatag
40

       Sequence 2616
       VTKASKDFVVNINICPPNRSWIAFVISTPGTIGMTAAMIIVAKSPLGTNACPITEIIPIT
       KEPMNSGINCACSFRSNISTTKYPIVPANTAVMISTRRPPCR*


45     Sequence 2617
       Contig_0784_pos_4022_5011,
       is similar to (with p-value 0.0e+00)
       >pir:pir|S39743|S39743 hypothetical protein - Bacillus subti
       lis
50     atggctgatttattatctgtattacaagacaaattatccgggaaaaatgtaaaaatagta
       ttacctgaaggtgaagatgaacgagtgctcattgctgcgactcagctacaaaaaaactgac
       tatgtttcacctatcgttctaggaaacgaagataatattaaatctcttgcttctaaacac
       gctttagatttaactcaaattgaaatcatagatccagcaacgagtgaacttaaagatgag
       cttgtagatgcttttgttgaaagacgtaaaggtaaggcaactaaagaacaagcagttgaa
55     ttattagataatgtaaattatttcggaacaatgcttgtgtatactggaaaggctgaaggt
       ttagtgagtggtgctgcacattctactggagatacagtcagaccagcattacaaattatc
       aaaactaaacctggtgtatctagaacatctggtattttcttatgattaaaggcgacgaa
       caatatattttttggagattgtgcgattaatccagaattagatgctcaaggacttgctgaa
       attgcagtagagagtgctaaatcagcacaaagctttggaatggaccctaaagtagctatg
```

```
ttaagcttttctacaaaaggttctgctaaatcggatgatgttactaaagtgcaagaagca
ttgaagttagctcaagaaaaagctgaagcagatcaattagatcatgtagttattgatgga
gaattccaatttgacgctgctattgttcctagcgtagcagagaagaaagcacctggtgca
aaaattcaaggtgatgcaaatgtattggttttccctagtctagaagcaggtaatattggt
tataagattgctcaacgtttaggtggatacgatgcagtaggaccagtcctacaaggatta
aactctccagtcaatgatttatctcgtggttgctcaactgaagacgtttataacttatct
attattacagctgctcaagctttacaataa
```

Sequence 2618
MADLLSVLQDKLSGKNVKIVLPEGEDERVLIAATQLQKTDYVSPIVLGNEDNIKSLASKH
ALDLTQIEIIDPATSELKDELVDAFVERRKGKATKEQAVELLDNVNYFGTMLVYTGKAEG
LVSGAAHSTGDTVRPALQIIKTKPGVSRTSGIFFMIKGDEQYIFGDCAINPELDAQGLAE
IAVESAKSAQSFGMDPKVAMLSFSTKGSAKSDDVTKVQEALKLAQEKAEADQLDHVVIDG
EFQFDAAIVPSVAEKKAPGAKIQGDANVLVFPSLEAGNIGYKIAQRLGGYDAVGPVLQGL
NSPVNDLSRGCSTEDVYNLSIITAAQALQ*

Sequence 2619
Contig_0784_pos_5083_5850,
is similar to (with p-value 5.0e-66)
>sp:sp|P39648|YWFL_BACSU HYPOTHETICAL 31.4 KD PROTEIN IN PTA
 3'REGION. >pir:pir|S39745|S39745 hypothetical protein - Bac
 illus subtilis >gp:gp|X73124|BSGENR_91 B.subtilis genomic re
 gion (325 to 333). NID: g413923. >gp:gp|Z99123|BSUB0020_60 B
 acillus subtilis complete genome (section 20 of 21): from 37
 98401 to 4010550. NID: g2636240.
```
atgcaatcttttgcgtttgatgacacttttccgaaagcgttggtaaagatttatcttgt
aatgtagtacgaacgtggatacatcaacacaccgtgattttgggcattcatgattcgcgt
ttaccattttttaagtgatggtattcgttttcttacagatgaacaaggatataatgcaatt
gttaggaattctggtggcttgggtgtcgtattagatcaaggaattttaaacatatctttg
atttttaaaggacaaaccgaaacgactattgatgaagcctttacagtgatgtatttattg
attaataaaatgtttgaggatgaagatgttagtatcgatactaaagaaattgagcaatcg
tattgcccaggaaaatttgatttaagtattaatgataagaaatttgccgggatttcgcag
cgacgagtacgtggtggtatcgcagtgcaaatatacttatgtattgaaggttctggctca
gaacgggcattaatgatgcaacagttttatcaacgtgcgcttaaaggggagactactaaa
tttcactatccagacatagatccctcatgtatggcatctttagaaacccttttaaataga
gaaattaaagtgcaagatgttatgtttttattattatatgcactaaaagatttaggggca
aacttaaatatggatcctattacagaagacgagtggacacgttacgaagggtattatgat
aagatgttagaacgcaatgcgaaaatgaatgaaaaattagatttttag
```

Sequence 2620
MQSFAFDDTFSESVGKDLSCNVVRTWIHQHTVILGIHDSRLPFLSDGIRFLTDEQGYNAI
VRNSGGLGVVLDQGILNISLIFKGQTETTIDEAFTVMYLLINKMFEDEDVSIDTKEIEQS
YCPGKFDLSINDKKFAGISQRRVRGGIAVQIYLCIEGSGSERALMMQQFYQRALKGETTK
FHYPDIDPSCMASLETLLNREIKVQDVMFLLLYALKDLGANLNMDPITEDEWTRYEGYYD
KMLERNAKMNEKLDF*

Sequence 2621
Contig_0784_pos_3890_3102,
is similar to (with p-value 9.0e-94)
>sp:sp|P39645|YWFI_BACSU HYPOTHETICAL 29.5 KD PROTEIN IN ROC
 C-PTA INTERGENIC REGION. >pir:pir|S39742|S39742 hypothetical
 protein - Bacillus subtilis >gp:gp|X73124|BSGENR_88 B.subti
 lis genomic region (325 to 333). NID: g413923. >gp:gp|Z99123
 |BSUB0020_63 Bacillus subtilis complete genome (section 20 o
 f 21): from 3798401 to 4010550. NID: g2636240.
```
atgaatccaataaaaattttgaaaaaggagcgaataataatgagtgaagcagcagaaact
ttagatggttggtatagcttacatttatttatgcagtagactggacaacttttcgttta
attgctgaagatgatcgtgaagcaatgattactgaattggaaacatttattaaagataaa
acagttgctagagaatcacatcaaggtgatcatgcaatttataacattacaggtcaaaaa
```

```
      gcggaccttttactatggtttttacgtccagaaatgaaagagttaaatcaaattgaaaat
      gagtttaataaattacgtatcgcagactatctcattccaacttattcctatgtgtcagtg
      atagaattaagtaattatttagcaggcaaatctgatgaggatccttatgaaaatccccac
      gttaaggcacgattataccctgaattaccacattctgaatatatatgtttctatccaatg
   5  gataaacgacgcaatgaaacttataactggtatatgttacctatcgaagaccgtaaaact
      ttaatgtataaccatgggatgataggtcgtaaatatgctggtaaaatcaaacagtttatt
      acaggttcagtaggttttgatgactatgagtggggtgttacattatttttcaaatgatgta
      cttcaattcaaaaaaattgtctatgaaatgcgtttttgatgaaacgactgctcgttatggc
      gaatttggtagtttctatattggtcacattctaaacatcgaagacttcaaacaatttttt
  10  agtatataa
```

Sequence 2622
MNPIKILKKERIIMSEAAETLDGWYSLHLFYAVDWTTFRLIAEDDREAMITELETFIKDK
TVARESHQGDHAIYNITGQKADLLLWFLRPEMKELNQIENEFNKLRIADYLIPTYSYVSV

  15 IELSNYLAGKSDEDPYENPHVKARLYPELPHSEYICFYPMDKRRNETYNWYMLPIEDRKT
LMYNHGMIGRKYAGKIKQFITGSVGFDDYEWGVTLFSNDVLQFKKIVYEMRFDETTARYG
EFGSFYIGHILNIEDFKQFFSI*

Sequence 2623
  20 Contig_0784_pos_2495_1845,
putative peptide of unknown function
```
      atggcacgtattgctacaaaattgggctatcctgaaagtaatagtttcgtgactaatact
      gtaattgaatttgttttacataatgaagcatatcctcggttgtatagaattaaaactcga
      gatacgaacttaataaaaatttctcaagctaatgaaatctcacgtcaaattacaaatggc
  25  acaatgacgcttgaagaagctaagtatcaattagaggaaatatatgttgctaaaagagat
      agcagtctcccttttaaaggaattgccgcagcaattatcgctacgagcttcctctatcta
      cagggaggtcgtctggttgatatcatcacagctgtattagctggaacgattggatactta
      gtagtagaaatattagatcgaaagctacacgcacaatttattccagaattcataggttct
      ttggtaataggtattatttctgtaattggacatgcatttgttcctagcggagatttagct
  30  acaattatcattgcagcggtcatgccgattgtacctggagtactgattacaaatgctatc
      caagatctattcggaggacatatgttaatgtttactacaaaatctttagaagctttagtc
      accgccttggtataggcgctggtgtaagttcaatattaattttagtctag
```

Sequence 2624
  35 MARIATKLGYPESNSFVTNTVIEFVLHNEAYPRLYRIKTRDTNLIKISQANEISRQITNG
TMTLEEAKYQLEEIYVAKRDSSLPFKGIAAAIIATSFLYLQGGRLVDIITAVLAGTIGYL
VVFILDRKLHAQFIPEFIGSLVIGIISVIGHAFVPSGDLATIIIAAVMPIVPGVLITNAI
QDLFGGHMLMFTTKSLEALVTAFGIGAGVSSILILV*

  40 Sequence 2625
Contig_0786_pos_2219_1125,
is similar to (with p-value 1.0e-53)
>sp:sp|P39604|YWCF_BACSU HYPOTHETICAL 43.3 KD PROTEIN IN QOX
D-VPR INTERGENIC REGION. >pir:pir|S39697|S39697 hypothetical
  45   protein - Bacillus subtilis >gp:gp|X73124|BSGENR_43 B.subti
lis genomic region (325 to 333). NID: g413923. >gp:gp|Z99123
|BSUB0020_107 Bacillus subtilis complete genome (section 20
of 21): from 3798401 to 4010550. NID: g2636240.
```
      atgggtggcggacaatatagtgcaaatttcagtattagacaaattatttattatattttc
  50  ggagctattattgcatttttaattatgataatttcaccgaaaaaaattaaaaataatact
      tatattttatacagtatattttgcgttctattaataggttacttattttacctgaaact
      tcaatcaccccaattattaatggtgctaaaagttggtacagtttcggtcctataagcatc
      caaccttccgaattcatgaaaattatacttatacttgctttggctaaaacgatatctaaa
      cataaccaatttacatttaataagtcttttcagtctgatttaatgttattttttaaaatt
  55  ttaggtgtatccattatacctatggcattaattctattgcaaatgacctaggtactact
      ttggtgttatgtgcaattatagctggcgtcatgttagtaagtggaataacatggaggata
      ttgjcccctctttttattgttgcatttgtaagtggttctagtattatattagctattatt
      tataaccatccttaatagaaaacctattaggaataaaaatgtatcaaatgggacgtatc
      aattcttggttagatccctattcatacagtagtggagatggatatcacttaacagaatct
```

```
     ttaaaagctattggttcaggtcaattattaggtaaaggttataaccatggcgaagtttat
     attcctgagaatcataccgactttattttttcagtgattggagaagaaatgggctttata
     ggttcagtattattgatattacttttcttattcttaatatttcaccttatacggttagct
     agtaaaattgatagtcagtttaacaaagtatttatcataggatatgtatcgttgattgtt
  5  tttcacgtgttacaaaatatcggcatgacggttcaattattaccgattacaggtatacca
     cttccgtttattagttacggtggaagttctttatggagtttaatgactggtataggagta
     gttctttcaatttattatcatgaaccccaaagatatgaaataaccacattatctaaaaaa
     tctaatacaatttaa
```

10 Sequence 2626
     MGGGQYSANFSIRQIIYYIFGAIIAFLIMIISPKKIKNNTYILYSIFCVLLIGLLILPET
     SITPIINGAKSWYSFGPISIQPSEFMKIILILALAKTISKHNQFTFNKSFQSDLMLFFKI
     LGVSIIPMALILLQNDLGTTLVLCAIIAGVMLVSGITWRILAPLFIVAFVSGSSIILAII
     YKPSLIENLLGIKMYQMGRINSWLDPYSYSSGDGYHLTESLKAIGSGQLLGKGYNHGEVY
15 IPENHTDFIFSVIGEEMGFIGSVLLILLFLFLIFHLIRLASKIDSQFNKVFIIGYVSLIV
     FHVLQNIGMTVQLLPITGIPLPFISYGGSSLWSLMTGIGVVLSIYYHEPQRYEITTLSKK
     SNTI*

     Sequence 2627
20 Contig_0787_pos_1817_2143,
     putative peptide of unknown function
     atgattttctttatcattaactttattcttatttcatatttgtacgataaacaatatgta
     cctttcaagcaattaccggtataagcttgtttattgcttagttatatttccaataaca
     ctcattttatatgtgcgtattgccaaaaaaattatctatacagtaataagtatgaaatg
25 agaactggaataatcattggtattattgctttaattctagtaattatgcaagggtttcac
     tttaactgggctatagatttttttaataatgttgtatggtggtcatttaaaagtaccatc
     tcaattgatagacgaatcgcaacatag

     Sequence 2628
30 MIFFIINFILISYLYDKQYVPFQAITGISLFYCLVIFPITLILYVRIAKKNYLYSNKYEM
     RTGIIIGIIALILVIMQGFHFNWAIDFFNNVVWWSFKSTISIDRRIAT*

     Sequence 2629
     Contig_0787_pos_6394_6038,
35 putative peptide of unknown function
     atgccagaaaatattgtaagaacaaaagggattgtatggttagcgcagtataatgatgta
     gcgtgtttgttatcacaggctggttcatcttgtaatattcaccccgttacatactgggtg
     gcaacaatgagtgaaagtcaacagcaagctattttggaggcgcgtcaagatgtagtagaa
     gattgggatatcgaatatggagatcgtcaaacgcaatttgtaattattggtacggattta
40 gatcaagaaaaaatttcccgggaattagatgcatgcttaattcatagtagtgagattgat
     gaagattggcgattactagatagtccgtatcaatggacttatgatcgacgaatgtaa

     Sequence 2630
     MPENIVRTKGIVWLAQYNDVACLLSQAGSSCNIHPVTYWVATMSESQQQAILEARQDVVE
45 DWDIEYGDRQTQFVIIGTDLDQEKISRELDACLIHSSEIDEDWRLLDSPYQWTYDRRM*

     Sequence 2631
     Contig_0787_pos_4438_2969,
     is similar to (with p-value 0.0e+00)
50 >sp:sp|P39755|NDHF_BACSU NADH DEHYDROGENASE SUBUNIT 5 (EC 1.
     6.5.3) (NADH-UBIQUINONE OXIDOREDUCTASE CHAIN 5). >gp:gp|U283
     23|BSU28323_1 Bacillus subtilis NADH dehydrogenase subunit 5
     (ndhF) gene, complete cds. NID: g903586. >gp:gp|Z99104|BSUB
     0001_183 Bacillus subtilis complete genome (section 1 of 21)
55 : from 1 to 213080. NID: g2632267. >gp:gp|Z99105|BSUB0002_11
     Bacillus subtilis complete genome (section 2 of 21): from 1
     94651 to 415810. NID: g2632457. >gp:gp|AB006424|AB006424_9 B
     acillus subtilis genomic DNA, 70 kb region between 17 and 23
     degree. NID: g3599592.
```

atgtttt.ttattacacttgtcattgctatactgagtggattaatattttttgaatc.tcga
gttcccattcaatatattaaatttcatatatatttacttgttttacctatcattacggga
ttaagtggattaatattttttcggtgaaagggcgaatgttggaccttttgtagttgatcat
cttacttggttaatgatgacatttattttgactttaggctttatcattcaaaagttttct
atgcgttatttaattggcgacatgcattaccgtaaatattttccgtttttttacattaatt
actgcatttgcttcattggcatggttaagtggagacttaaggttaatgaccatgttttgg
ggtgcaacattatttgtgttaacacggctcattaaagttaacaaattatggaaggtgcct
agggaagcagcaagaatttcagcttggtcatttatattggcatggttgtcgttattgatt
gctgtcattttattgtatatcgctacaggagattggtatatttattcgaatatgtcagat
gataatgcaatcaattatggaatgcgtctctgtatcaatttacttattgtttttagctgtg
attattccggcggcacaatttccatttcaaggctggcttattgaatctgtagctgcgcct
acgccagtttcagctattatgcacgctggtattgttaatgctggtggcgttattcttaca
cgctttttctccggtatttaatgacgaaatagccatttcactgttattaattattgcaagt
atttcagtattgttgggttctggaatcagtcttgtgcatgttgattacaagagacaactt
gtaggttctacgataagtcaaatgggttttatgttagtacaatgtgcgcttggggcatat
tctgcggcgatagtacatttaatattgcatggtgtgtgtttaaagcgacattatttctacaa
tcgggttctgttgttaaaagatttaacattcctacgcctccatctgttaaaaaatcatat
ggctggcttgtatttggtcgtctactagctattcttatagcgataatattttggttgaat
agtgatagacatgcatatgatgtattaagcgctcttatattagcttggtcgttaatggtg
tc.tqgaatcaattagttgcttttagtcatggactcatcggaagagtcatcggagt.tgt
atgattattgttgtagtaattgtttatattattacgcatcattatttcttcacgar.atta
agtaacgttgatattcatattgtttcaccaccactcataagtattattctatcgattgct
attatagtttttcggcagtatgttaagcatatgggtatcacggcgaagagaatcaaaggca
tttgcgaagttatacttgtggcttattaaagtaggagaggctaaaacccaatctatagaa
agtcatccatcatatttaaaacgattttag

Sequence 2632
MFFITLVIAILSGLIFLNHRVPIQYIKFHIYLLVLPIITGLSGLIFFGERANVGPFVVDH
LTWLMMTFILTLGFIIQKFSMRYLIGDMHYRKYFPFFTLITAFASLAWLSGDLRLMTMFW
GATLFVLTRLIKVNKLWKVPREAARISAWSFILAWLSLLIAVILLYIATGDWYIYSNMSD
DNAINYGMRLCINLLIVLAVIIPAAQFPFQGWLIESVAAPTPVSAIMHAGIVNAGGVILT
RFSPVFNDEIAISLLLIIASISVLLGSGISLVHVDYKRQLVGSTISQMGFMLVQCALGAY
SAAIVHLILHGVFKATLFLQSGSVVKRFNIPTPPSVKKSYGWLVFGRLLAILIAIIFWLN
SDRHAYDVLSALILAWSLMVSWNQLVAFSHGLIGRVIGVCMIIVVVIVYIITHHYFFTTL
SNVDIHIVSPPLISIILSIAIIVFGSMLSIWVSRRRESKAFAKLYLWLIKVGEAKTQSIE
SHPSYLKRF*

Sequence 2633
Contig_0787_pos_2951_2058,
putative peptide of unknown function
atgcttcaatcagatatcaatgaattagtcaatcaggctaaacgtgtaattacacttta
tcaccgatttcaacatttgctgcccgtaatccgtgggagggggctagaagatgcttcgttt
gatcaagtggcacgttggttaaaaagtgtgagggatgttgacatttatcctaatgcgtct
actattcacagagcgattagtaataaagaaatagatttaaaagtatttgaagaacggttg
gatgaaaatcgtgcgcattataataataggtcactatctgacagtgatatcaacacatat
attcaaagagcgaaaaatttaaaaacgattgaagaaggttactttaatacaaaagataac
gagaaactggaaaatgggtacaaactaattttaaggattataagaaaaaagaagatgtg
atagcgcaaagtgctagtgttttcacaaaggaaggtacacgacttattgatattttaaat
gctcatatgattaagtggtctaaattatatgttgatgactttcaatcaagttggactatg
ccaaaaagagaaaaaggattctatcatgcctggcaacgtttagttaaacatgatccatta
ttcacaaaaaaacaacgacttactttagcacatttgccaaatcaagcaaccgaagcaata
gagtacgcctttcaagaattaggagtaaaagaagaacatcgacaatcatatattgagagt
catttattatctttaccaggttgggcaggaatcatgtatcatcggtcacagacacaaagt
aatgatgcgtacttattaacagactatgttgcgattcgtctatcaattgagatggtactt
ttaaatgaccaccatacaacattattaaaaaaatctatagcccagttaaagtga

Sequence 2634
MLQSDINELVNQAKRVITPLSPISTFAARNPWEGLEDASFDQVARWLKSVRDVDIYPNAS
TIHRAISNKEIDLKVFEERLDENRAHYNNRSLSDSDINTYIQRAKNLKTIEEGYFNTKDN

EKLEKWVQTNFKDYKKKEDVIAQSASVFTKEGTRLIDILNAHMIKWSKLYVDDFQSSWTM
PKREKGFYHAWQRLVKHDPLFTKKQRLTLAHLPNQATEAIEYAFQELGVKEEHRQSYIES
HLLSLPGWAGIMYHRSQTQSNDAYLLTDYVAIRLSIEMVLLNDHHTTLLKKSIAQLK*

5   Sequence 2635
    Contig_0787_pos_0_876,
    is similar to (with p-value 0.0e+00)
    >pir:pir|A40585|A40585 DNA topoisomerase (ATP-hydrolyzing) (
    EC 5.99.1.3) chain B - Staphylococcus aureus >gp:gp|X71437|S
10  AGYRREC_2 S.aureus genes gyrB, gyrA and recF (partial). NID:
     g296393. >gp:gp|D10489|STAGYRABA_1 Staphylococcus aureus ge
    nes for DNA gyrase A and B, complete cds. NID: g540540.
    gtgatggtgaatacattgtcagatgtaaacaacacagataattatggtgctggacagata
    caagttttagaaggtctcgaagcggttcgtaaaagaccgggtatgtatattggttcaact
15  tcagaaagagggttgcaccatttagtatgggaaattgttgataatagtattgacgaggca
    ttagcaggttatgctagtcatattgaagttgtaattgagaaagacaattggattaaagtt
    actgacaatggccgtggtattcctgttgatattcaagaaaagatgggacgccctgctgtc
    gaagttatcttaactgtacttcacgctggaggtaaattcggaggtggcggatacaaagta
    tctggcggtcttcacggtgttggatcttcagttgttaatgcactctcacaagatcttgaa
20  gtttatgtacatcgtaatggcacgatttatcatcaagcctataaacaaggtgtgccacaa
    tttgatcttaaagaaattggcgatacagataaaacaggtacagctattcgattcaaagcc
    gataaagaaatctttacagagacaacagtttataactatgaaacacttcaaaagcgtata
    cgtgagcttgctttcttaaataaaggtattcaaattacttaaaagatgaaagagaagag
    gaagttagagaagactcatatcattatgaaggcgggattaaatcctatgtagatttatta
25  aatgagaataaagaacctcttcacgatgaacctatatatatccatcagtctaaagacgat
    attgaagtggaaattgcacttcaatataacagtggatatgcaaccaacttattaacgtat
    gcgaataatattcatacatacgagggtggtaTTCAA


    Sequence 2636
30  VMVNTLSDVNNTDNYGAGQIQVLEGLEAVRKRPGMYIGSTSERGLHHLVWEIVDNSIDEA
    LAGYASHIEVVIEKDNWIKVTDNGRGIPVDIQEKMGRPAVEVILTVLHAGGKFGGGGYKV
    SGGLGVGSSVVNALSQDLEVYVHRNGTIYHQAYKQGVPQFDLKEIGDTDKTGTARFKA
    DKEIFTETTVYNYETLQKRIRELAFLNKGIQITLKDEREEEVREDSYHYEGGIKSYVDLL
    NENKEPLHDEPIYIHQSKDDIEVEIALQYNSGYATNLLTYANNIHTYEGGIQ
35

    Sequence 2637
    Contig_0788_pos_5197_6411,
    is similar to (with p-value 3.0e-17)
    >pir:pir|I64093|I64093 ribosomal protein S14 (rpS14) homolog
40   - Haemophilus influenzae (strain Rd KW20) >gp:gp|U32762|U32
    762_5 Haemophilus influenzae Rd section 77 of 163 of the com
    plete genome. NID: g1573797.
    gtggatgatgtgacaaaatatggtccagttgatggagatccgatcacgtcaacggaagaa
    attccattcgacaagaaacgtgaattcaatcctgatttaaaaccaggtgaagagcgtgtt
45  aaacaaaaaggtgaaccaggaacaaaaacaattacaacaccaacaactaagaacccatta
    acaggggaaaaagttggcgaaggtgaaccaacagaaaaaataacaaaacaaccagtagat
    gaaatcacagaatatggtggcgaagaaatcaagccaggccataaggatgaatttgatcca
    aatgcaccgaaaggtagccaagaggacgttccaggtaaaccaggagttaaaaaccctgat
    acaggcgaagtagtcacaccaccagtggatgatgtgacaaaatatggtccagttgatgga
50  gatccgatcacgtcaacggaagaaattccattcgacaagaaacgtgaattcaatcctgat
    ttaaaaccaggtaaagagcgcgttaaacagaaaggtgaaccaggaacaaaaacaattaca
    acaccaacaactaagaacccattaacaggggaaaaagttggcgaaggtgaaccaacagaa
    aaagtaacaaaacaaccagtagatgaaatcacagaatatggtggcgaagaaatcaagcca
    ggccataaggatgaatttgatccaaatgcaccgaaaggtagccaagaggacgttccaggt
55  aaaccaggagttaaaaatcctgatacaggcgaagtagttactccaccagtggatgatgtg
    acaaaatatggtccagttgatggagatccgattacgtcaacggaagaaattccgtttgat
    aaaaaacgcgaatttgatccaaacttagcgccaggtacagagaaagtcgttcaaaaaggt
    gaaccaggaacaaaaacaattacaacaccaacaactaagaaccctatggcgaagaaatct
    aaaatagcaaaagaacaaaaaagacaagaattagtaaataaatattacgagttacgtaaa

```
gaattaaaagcaaaaggggactatgaagcattaagaaagttgccaagagattcatctcca
actagattaactagaagatgtaaagtaactggtagacctagaggtgtgttacgtaaattt
gaaatgtctagaattgcatttagagaacatgcgcataaaggtcaaattccaggtgtaaaa
aaatctagttggtaa
```
5

Sequence 2638
VDDVTKYGPVDGDPITSTEEIPFDKKREFNPDLKPGEERVKQKGEPGTKTITTPTTKNPL
TGEKVGEGEPTEKITKQPVDEITEYGGEEIKPGHKDEFDPNAPKGSQEDVPGKPGVKNPD
TGEVVTPPVDDVTKYGPVDGDPITSTEEIPFDKKREFNPDLKPGKERVKQKGEPGTKTIT
10      TPTTKNPLTGEKVGEGEPTEKVTKQPVDEITEYGGEEIKPGHKDEFDPNAPKGSQEDVPG
KPGVKNPDTGEVVTPPVDDVTKYGPVDGDPITSTEEIPFDKKREFDPNLAPGTEKVVQKG
EPGTKTITTPTTKNPMAKKSKIAKEQKRQELVNKYYELRKELKAKGDYEALRKLPRDSSP
TRLTRRCKVTGRPRGVLRKFEMSRIAFREHAHKGQIPGVKKSSW*

15      Sequence 2639
Contig_0788_pos_5829_4909,
putative peptide of unknown function
```
gtgatttcatctactggttgtttttgttactttttctgttggttcaccttcgccaactttt
tcccctgttaatgggttcttagttgttggtgttgtaattgttttttgttcctggttcacct
20   ttctgtttaacgcgctctttacctggtttttaaatcaggattgaattcacgtttcttgtcg
aatggaatttcttccgttgacgtgatcggatctccatcaactggaccatattttgtcaca
tcatccactggtggtgtgactacttcgcctgtatcagggtttttaactcctggtttacct
ggaacgtcctcttggctacctttcggtgcatttggatcaaattcatccttatggcctggc
ttgatttcttcgccaccatattctgtgatttcatctactggttgttttgttatttcttct
25   gttggttcaccttcgccaactttttcccctgttaatgggttcttagttgttggtgttgta
attgttttgttcctggttcaccttttgtttaacacgctcttcacctggttttaaatca
ggattgaattcacgtttcttgtcgaatggaatttcttccgttgacgtgatcggatctcca
tcaactggaccatattttgtcacatcatccacaggtggagtaactacttcgcctgtatca
ggatttttaaccccccggcttacctggttgcgtgtttgactacctttcggtgcatttgg
30   atcaaattcatccttatggcctggcttgatttcttcgccaccataatgaacgatttcatc
cactggttgtttttgttattttttctgttggttcaccttcgccaactttttctcctgtatt
aggattgacataagttggtgttgttgttgtttcaattcctggttcacctttttggactac
tttttctgtacctggggctaa
```

35      Sequence 2640
VISSTGCFVTFSVGSPSPTFSPVNGFLVVGVVIVFVPGSPFCLTRSLPGFKSGLNSRFLS
NGISSVDVIGSPSTGPYFVTSSTGGVTTSPVSGFLTPGLPGTSSWLPFGAFGSNSSLWPG
LISSPPYSVISSTGCFVIFSVGSPSPTFSPVNGFLVVGVVIVFVPGSPFCLTRSSPGFKS
GLNSRFLSNGISSVDVIGSPSTGPYFVTSSTGGVTTSPVSGFLTPRLTWLRCLTTFRCIW
40      IKFILMAWLDFFATIMNDFIHWLFCYFFCWFTFANFFSCIRIDISWCCCCFNSWFTFLDY
FFCTWG*

Sequence 2641
Contig_0788_pos_3383_2700,
45      is similar to (with p-value 2.0e-16)
>sp:sp|P54176|HLY3_BACCE HEMOLYSIN III (HLY-III). >pir:pir|S
59967|S59967 hemolysin III - Bacillus cereus >pir:pir|S52296
|S52296 hemolytic factor - Bacillus cereus >gp:gp|X84058|BCH
MLYSN_1 B.cereus gene for hemolytic factor. NID: g662879.
50      atgtctcaatcatctaagcgaaaaaacgattctgtggtagaaacttttaaggacatcatt
cctttaacatttggagaagagattggtaatgcagcatcacatggtgctgctgcattactt
acattatttattcttccatatgcagcagtgcatagtttttaacaatggtggcacattagag
tcaattagtgtgtcagtttatgtgattagtattttttatgatgttcatatcatcaaccatt
tatcattcaatgcaaaataatacgtctcataaatatatattaaggattattgaccatagt
55      atgatttatgtggctatttcaggaacatacacacctgttttgttaagtgttgtcggcggt
tggttaggttggctcgtgacaatattattatggggaacgacattatgggggattttgtat
aaatcaatagcaactaaagtcaatcatagattaagtctcattgtttatttggtgatggga
tgggtaggtatcatattttacctattattattatgcgaacatcatggtggtttatttttc
tttatatttcttggtgggttatcttatactatcggagcatggttttatgcccaaaaaaat
```

aaaccttattttcatatgatttggcatatatttattgttcttgcttctttcttacatatg
ataggcattttttattttatgtga

Sequence 2642

5   MSQSSKRKNDSVVETFKDIIPLTFGEEIGNAASHGAAALLTLFILPYAAVHSFNNGGTLE
SISVSVYVISIFMMFISSTIYHSMQNNTSHKYILRIIDHSMIYVAISGTYTPVLLSVVGG
WLGWLVTILLWGTTLWGILYKSIATKVNHRLSLIVYLVMGWVGIIFLPIIIMRTSWWFIF
FIFLGGLSYTIGAWFYAQKNKPYFHMIWHIFIVLASFLHMIGIFYFM*

10   Sequence 2643
Contig_0788_pos_2461_1124,
is similar to (with p-value 7.0e-20)
>gp:gp|D50098|D50098_1 Bacillus subtilis macronuclear DNA fo
r multidrug transporter, complete cds. NID: g1856976.
15   atgastc.taagtctatcattactgtaatggcactcatactaataatgtttatgg·agct
atagaaacatctattatttcattagcattaccaacaataaaaaatagtttgaatgccggt
aatctagtttcattagtatttaccgtatattttattgctttagtcatagctaaccctatc
gttggtgaacttatgtctagatttaaaattatttacattgctgttgtaggggtattattg
tttgccttaggtagtttaatgtcgggattaagtcagacgtttacttttttaattatctct
20   cgaacagtacaaggttttggagcaggagttatgatgtcactctcacaaatagttcctaag
ttggcttttgaaattcctttgagatataaaattatgggtatagttggaagtgtttgggga
atttcgagtattattggcccattattaggtggtgcgatttttagagtttgcttcatggcat
tggctattctatatcaatattcctattgctatagtggcaataatacttgtacttatgact
tttcatttcctgatgagacacaagtacaacagagtcgttttgatataaaaggattgatt
25   atcttttatatctttatagctttattaatgtttggtttactcaaccaacatcatattatt
tttaacatattctcaattattttagctttagctgtttatggctactatttaaaatagaa
aatagtatcgaacaaccatttcttccaacaaagaatttaacatatcaatagttctagtt
tttataacggatttacttattgcgataacactgatgggatataatttatatataccagta
tatttacaagaaaaacttagtttatcacctttacaaagtggatttgtaatattcccgttg
30   tctgttgcttggattacgcttaatttcaatttagctaaaatagaagcgcattttactaga
aaaacattatatatttgctcattttttgttttattagttagtagtctgatgataatgttt
ggcttaaaactcccattgcttattgcttttgcggttgttttgcaggtttaagttttggt
tatatttatacaaaagatagtgttattgtccaagaggaaacttctccaaaaaatatgaaa
aagatgatgtcattttatgcattgacaaaaaatttaggttcgtcagtcggatctacgatt
35   atgggctatatgtatgcactaaatgttggtttatttggttctaatttacacaatgratta
ggattagtcttaataattgcagtatgtttaattgtaatgtggatgacattatataaaagc
aatactattcaatcttag

Sequence 2644
40   MNLKSIITVMALILIMFMAAIETSIISLALPTIKNSLNAGNLVSLVFTVYFIALVIANPI
VGELMSRFKIIYIAVVGVLLFALGSLMSGLSQTFTFLIISRTVQGFGAGVMMSLSQIVPK
LAFEIPLRYKIMGIVGSVWGISSIIGPLLGGAILEFASWHWLFYINIPIAIVAIILVLMT
FHFPDETQVQQSRFDIKGLIIFYIFIALLMFGLLNQHHIIFNIFSIILALAVLWLLFKIE
NSIEQPFLPTKEFNISIVLVFITDLLIAITLMGYNLYIPVYLQEKLSLSPLQSGFVIFPL
45   SVAWITLNFNLAKIEAHFTRKTLYICSFFVLLVSSLMIMFGLKLPLLIAFAVVFAGLSFG
YIYTKDSVIVQEETSPKNMKKMMSFYALTKNLGSSVGSTIMGYMYALNVGLFGSNLHNVL
GLVLIIAVCLIVMWMTLYKSNTIQS*

Sequence 2645
50   Contig_0795_pos_9882_9001,
putative peptide of unknown function
atgccaagttcaatgttgcatgagcatggatggagcggtatcaatttcaattcaccattt
gcagatatggcaatttttattaggtcttaactggttagcaatattactttatatggaagca
gtl.jt.jt.caccgtttggtactggagtttcttttgttgccgttactggacgtgtgtt.acgc
55   gctatgçaagaaaatgggcatattcctaaattcttaggtaaaattaataaaaagtataat
atcccacgtgttgccattgcatttaatgcaattatcagcatgattatggtgacattgttc
cgtgactggggtacactagctgcggttatttctactgcaacattagttgcatatttaact
ggtccaactacggttatttcattacgtaaaatggcaccaaaaatgactcgtccatttaaa
gctaatattttaaaatttatggcacctttatcctttgttttagcatcattagctatctat

```
tgggcaatgtggccaacaacagcagaagttattttaattattattttaggtttacctatt
tatttcttctatgaatataaaatgaattggaagaatactaaaaaacaaattggcggaagc
ttatggattattatctatcttattgttctcgcattcttatcatttattggaagtaaagag
ttcaaaggcttaaattggattcactatccatgggatttcttagtcattgtaatcgttgct
```
5   `ttaatcttctatcaactaggtacgacaagttactttgaaagtatttatttcaagcgtgcg`
```
aacaagttgaataagaaaatgggcgataagttgcgtaaaacacgcaaaaaagcgcgtcat
aaagatrggaaagaacgcgatcgacaagagcaaaatcaataa
```

Sequence 2646
10   MPSSMLHEHGWSGINFNSPFADMAILLGLNWLAILLYMEAVVSPFGTGVSFVAVTGRVLR
AMEENGHIPKFLGKINKKYNIPRVAIAFNAIISMIMVTLFRDWGTLAAVISTATLVAYLT
GPTTVISLRKMAPKMTRPFKANILKFMAPLSFVLASLAIYWAMWPTTAEVILIIILGLPI
YFFYEYKMNWKNTKKQIGGSLWIIIYLIVLAFLSFIGSKEFKGLNWIHYPWDFLVIVIVA
LIFYQLGTTSYFESIYFKRANKLNKKMGDKLRKTRKKARHKDWKERDRQEQNQ*

15

Sequence 2647
Contig_0795_pos_8696_7329,
is similar to (with p-value 2.0e-47)
>gp:gp|U46134|BSU46134_3 Bacillus subtilis putative orf1 unk
20   nown protein, putative transcriptional regulator (slr), and
intracellular esterase B (estB) genes, complete cds. NID: g1
762123. >gp:gp|Z99121|BSUB0018_125 Bacillus subtilis complet
e genome (section 18 of 21): from 3399551 to 3609060. NID: g
2635827. >gp:gp|Z71928|BSYVEFGNS_1 B.subtilis pnbA, sigL, yv
25   e[J,K,L,M,N,O,P,Q,R,S,T] and yvf[A,B,C,D,E,F,G,H] genes. NID
: g1495276. >gp:gp|Z94043|BSZ94043_47 B.subtilis genomic DNA
fragment (88 kb). NID: g1945641.

```
gtgtgttcaaacatggttcaagtcaagataggtaactgtaccatcaatggcttacacaaa
aagaatattgatgtattcttaggtatcccgtatgccaaaccattcaataagatatctcga
```
30   `ttccaacattcaaagtttatggaactaagcaaaccaatgattgatgcaactcatattcaa`
```
tccatcccaccacaaccctacaattcacttgaagacttttttcgatgacagattcatcg
tttaattcttttaaacaaaatgattattgcctgtttttaaatatttggaaaccatcgtcc
aatcaaaatcatttacctgtagtgatttactttatggtggtagttttcttcaaggacat
ggcacagctgaactatattgtcctgaacacatagtagaacaagaaaatataataagtagtt
```
35   `acttttaattatcgcttaggtgcactcggctacctagattggtcttatttttaatcaacat`
```
ttgaactataataatgggatttctgatcaaattaacgtttttaagatgggtacatcaatat
atcgaacattttggcggtgattcaaataacgtgacactaatgggtcaatctgcaggtagt
atgagtatcatgacattaatgcaaatgcccgaacttgatgattattatcataaagtgatg
ttattaagtggaacgttaactactgatacaccactcaatgcacatactaaagtacaacat
```
40   `ttttcacaactcatgaggcattattttcctaataaaacacttaagacacttaccagtgat`
```
gacattttatatctaatggagtctcaaaaaatagagcgtggaagatctcgtggacttgat
ttgatttatcaacctattaaagatcatcatatgtcacgatccattaaaaaatttcccaaa
ccgacattcatgagttatacacacgatgaaggtgatatttatattgaagacgcaacacgc
accttaccttctgaacgttttattcacttgatgtctcaatatggtacacacgtcgaaaaa
```
45   `aatgatgccctcacaatgaaacaacaaagaaatttaataacagagtattgttttgttcgt`
```
ccaatttatctattttttaaatcaaatgaatagttgcgacacttggctagcacgttttgat
tggcaccaacccccatacctcctactttaaaagtgcatatcatatattggatttagtattt
tggtttggtcacctctctattttgactaaaaatcattattctataactcaacatgatatg
aatttaagtagtaacatgatatctgacttagcttattttgcccgaaaaggtaagatgcca
```
50   `tggaaatgttatgaacctcaacatcaagcgttacatatctatggataa`

Sequence 2648
VCSNMVQVKIGNCTINGLHKKNIDVFLGIPYAKPFNKISRFQHSKFMELSKPMIDATHIQ
SIPPQPYNSLEDFFSMTDSSFNSFKQNDYCLFLNIWKPSSNQNHLPVVIYFYGGSFLQGH
55   GTAELYCPEHIVEQENIIVVTFNYRLGALGYLDWSYFNQHLNYNNGISDQINVLRWVHQY
IEHFGGDSNNVTLMGQSAGSMSIMTLMQMPELDDYYHKVMLLSGTLTTDTPLNAHTKVQH
FSQLMRHYFPNKTLKTLTSDDILYLMESQKIERGRSRGLDLIYQPIKDHHMSRSIKKFPK
PTFMSYTHDEGDIYIEDATRTLPSERFIHLMSQYGTHVEKNDALTMKQQRNLITEYCFVR
PIYLFLNQMNSCDTWLARFDWHQPHTSYFKSAYHILDLVFWFGHLSILTKNHYSITQHDM

650
```

NLSSNMISDLAYFARKGKMPWKCYEPQHQALHIYG*

Sequence 2649
Contig_0795_pos_2619_2143,
is similar to (with p-value 4.0e-68)
>gp:gp|Z99108|BSUB0005_72 Bacillus subtilis complete genome
(section 5 of 21): from 802821 to 1011250. NID: g2633055. >g
p:gp|D78509|D78509_8 Bacillus subtilis YfjG-YfjR genes, comp
lete cds. NID: g2780390.
atggaagacgtgacagatattgtctttcggcatgttgtcagtgaagctgcgagaccagat
gtattttttactgaatttaccaatactgagagttactgtcaccctgaaggtattcatagt
gtgcgcggacgcttaacttttagtgacgacgaacaaccaatggtagcgcacatctggggc
gataaaccagaacaattccgagaaatgagtatcggcttagcggatatgggttttaaaggt
atagatttaaatatgggttgccctgtcgcaaacgttgcgaaaaaaggtaaaggatccggc
ttaattctacgacctgaaacggcagccgaaatcattcaagcttctaaagcaggtggtcta
ccggtcagtgtaaaaacacgtttaggttattacgatatcgatgaatggcgagactggtta
aaacacgtcttcgaacaagttaggtgcgcgctctggtttaatggcagagccaaatga

Sequence 2650
MEDVTDIVFRHVVSEAARPDVFFTEFTNTESYCHPEGIHSVRGRLTFSDDEQPMVAHIWG
DKPEQFREMSIGLADMGFKGIDLNMGCPVANVAKKGKGSGLILRPETAAEIIQASKAGGL
PVSVKTRLGYYDIDEWRDWLKHVFEQVRCALWFNGRAK*

Sequence 2651
Contig_0795_pos_2063_660,
is similar to (with p-value 0.0e+00)
>gp:gp|AF029225|AF029225_1 Staphylococcus carnosus NarG, Nar
H, NarJ, and NarI genes, complete cds. NID: g3929521.
atgactgcgacgccattatattcagatatcgttttacctgctgcaacttggtatgaaaaa
catgatttatcttctacagacatgcatccatttattcatccatttaacccagcgattgac
ccattatgggaatcgcgttcggactgggatatttataaaactctaagtaaagctgtttca
gaaatggcgaaagattatcttccaggtaaatttaaagatgtcgtaactacaccattagga
catgattcaaaacaagaaatttcaactgaatacgtattgtaaaagattggtctaaagga
gaaattgaaggtgtgccaggtaaaacaatgcctaatttttctatcgtagagcgagactat
acacaaatttacgataaaattcgttactgttggtccaaaactagaaaaagggaaaataggt
gctcatggtgtgagttatagcgttagtgaagagtacgaagaacttaaagtatagttgga
acttcgaatgatgataatactatttcagttaaaaatgatagaccgagaatagatacagcg
agaaaactagcagatgtcattttgaatatatcctctgctacaaacggcaaattatcacaa
aagtcatatgaagatttagaaaatcaaacaggtatggaacttaaagatatttctaaagaa
cgtgcttctgaaaagatatcattcttaaacattacttctcaaccaagagaagtgattcca
actgcagtattccctggctctaataaagatggaagacgctactcaccgtttacaactaat
gttgaacgtttagtgccatttagaacactaactggacgtcaaagttattatatagatcat
gaggtattccaacagtttggcgaaagtttaccggtatataaacctactttacctccaatg
gtatttggtgctcgtgataaaaaagttaaaggtggacaagatacattagtgcttcgatac
cttacacctcatggaaaatggaatattcattcaacttatcaagataatgaacgcatgttg
acgttgtttagaggtggaccagttgtatggatttcaaatgaagacgcagctgaccatggt
attaatgataacgactggttagaagtatacaacagaaacggagttgttactgccagagct
gtaacatctcatcgtatgcctagaggcacaatgtttatgtatcatgcacaagataaacat
atagagacacctggttctgaaattactgatactcgtggaggttctcataatgcacctact
cgtattcacttgaaacctactcaattagtaggaggatatgcacaaattagttatcacttt
aactattatggaccaattggaaatcaaagagatgagtatgtagctgttagaaaaatgaag
gaggtcaattggcttgaagattaa

Sequence 2652
MTATPLYSDIVLPAATWYEKHDLSSTDMHPFIHPFNPAIDPLWESRSDWDIYKTLSKAVS
EMAKDYLPGKFKDVVTTPLGHDSKQEISTEYGIVKDWSKGEIEGVPGKTMPNFSIVERDY
TQIYDKFVTVGPKLEKGKIGAHGVSYSVSEEYEELKSIVGTWNDDNTISVKNDRPHIDTA
RKVADVILNISSATNGKLSQKSYEDLENQTGMELKDISKERASEKISFLNITSQPREVIP
TAVFPGSNKDGRRYSPFTTNVERLVPFRTLTGRQSYYIDHEVFQQFGESLPVYKPTLPPM

VFGARDKKVKGGQDTLVLRYLTPHGKWNIHSTYQDNERMLTLFRGGPVVWISNEDAADHG
INDNDWLEVYNRNGVVTARAVTSHRMPRGTMFMYHAQDKHIETPGSEITDTRGGSHNAPT
RIHLKPTQLVGGYAQISYHFNYYGPIGNQRDEYVAVRKMKEVNWLED*

5   Sequence 2653
    Contig_0795_pos_0_646,
    is similar to (with p-value 0.0e+00)
    >gp:gp|AF029225|AF029225_2 Staphylococcus carnosus NarG, Nar
    H, NarJ, and NarI genes, complete cds. NID: g3929521.
10  atggtattgaatctagacaaatgtattggttgtcatacttgcagtgtgacatgtaaaaac
    acatggacaaatcgacctggtgcagaatatatgtggtttaataacgtagaaacaaaaccg
    ggtgtaggatatccaaaaagatgggaagaccaaggacaatataaaggtggttgggtgcta
    aataaaaaggaaagcttgaattaaaatctggtaacagatggtcaaaaattgctttaggt
    aaaatcttctataatccagacatgccactcattcaagattattatgaaccgtggacatat
15  aactatgaacacttaaccaatgctaaacaaggacagcactctcccgtggcgacagctcac
    tctttaatttcaggtgatagattgaatcttaaatgggggccaaactgggaagatgattta
    gctggaggtcacattacaggaccagaggatccaaatattcagaaaatagaagaagatatt
    aaattccaattcgatgagacatttatgatgtatttaccaagactatgtgaacactgttta
    aatccaagttgcgtagcatcttgtccatcaggagctatgtataaacgagatgaggatggt
20  atcgtactcgtcgatcaagaagcctgtcgaggttggagatactgta

    Sequence 2654
    MVLNLDKCIGCHTCSVTCKNTWTNRPGAEYMWFNNVETKPGVGYPKRWEDQGQYKGGWVL
    NKKGKLELKSGNRWSKIALGKIFYNPDMPLIQDYYEPWTYNYEHLTNAKQGQHSPVATAH
25  SLISGDRLNLKWGPNWEDDLAGGHITGPEDPNIQKIEEDIKFQFDETFMMYLPRLCEHCL
    NPSCVASCPSGAMYKRDEDGIVLVDQEACRGWRYCX

    Sequence 2655
    Contig_0796_pos_1625_2248,
30  putative peptide of unknown function
    atggattatgtatacacaatttataaaaatcctagatataacatcattcaaaaagataat
    cgctatttaatggtcgatttagaacaaaattggtattcatatttatgtccaatgctgaat
    tggtttatacctattaaattcacagaattaacttatcaagaattcaataatataaacata
    tttcataatggaggacaaaagagtcatggtatgatggctgctggcgttggcgtcactatc
35  agtgtgctattaagaagtcttgtgggttatatagatattaatattagtcgaatttggata
    gtttttatgttttttaattggatttgttgctgtgatcacacttcgtttatctataagaaag
    aagttaaatcatccagcatttaataaaaagagtaaacaaaaagtaatattgataccatca
    tttaaaaatatgatattggtggtgtttttgctattttatgatgctgttttttctcaattgca
    ccttttcaaatgattttttgaggaaaagaaaaacatcttaggatatatactttgggtaggt
40  gtattatttatatttactactttgaatatggcttcaatttctgatagaaaagtacatgcc
    aaaattaaaaatataagaagatag

    Sequence 2656
    MDYVYTIYKNPRYNIIQKDNRYLMVDLEQNWYSYLCPMLNWFIPIKFTELTYQEFNNINI
45  FHNGGQKSHGMMAAGVGVTISVLLRSLVGYIDINISRIWIVFMFLIGFVAVITLRLSIRK
    KLNHPAFNKKSKQKVILIPSFKNMILVVFCYFMMLFFSIAPFQMIFEEKKNILGYILWVG
    VLFIFTTLNMASISDRKVHAKIKNIRR*

    Sequence 2657
50  Contig_0796_pos_4131_4508,
    putative peptide of unknown function
    atgcaattttattatagtaattgtgaacagaatgcatcaaacttattagtcatcaacgta
    caacctaatgaaggattttctttatgtgtgaatggtaagaaaagtaatcaaaataatgaa
    atgcaaaaagtgaagctttcttatactatgccgattaaagataaaatgaacacagttgat
55  gcatatgaaatcttatttacgatacattaattggagaacaaacaaaatttacgcattgg
    gaagaattaaaaattcttggaaatttattgatgatattgaaaatgtatggaacaagaat
    agccacagtttcctaattatgcctttggatgctatgggcctaaagaaagtgaaaaattac
    ttagtgaagacggattga

Sequence 2658
MQFYYSNCEQNASNLLVINVQPNEGFSLCVNGKKSNQNNEMQKVKLSYTMPIKDKMNTVD
AYENLIYDTLIGEQTKFTHWEELKILGNLLMILKMYGNKNSHSFLIMPLDAMGLKKVKNY
LVKTD*

5

Sequence 2659
Contig_0796_pos_4627_0,
is similar to (with p-value 6.0e-53)
>gp:gp|L76359|STMDRRC_1 Streptomyces peucetius daunorubicin

10  resistance protein (drrC) gene, complete cds. NID: gl196906.
atggattttattaatattacaggtgcttcacaaaataacttgaaaaacatagatgtaaat
atcccaaaacacttagtaacggtatttacaggtcgttctggttcagggaaatcatcttta
gtgtttaatactgttgctgcggagtctgaacagctactaaatgaaagttattctagttat
attcaatttcatttaaatcaacaacccagaccgaaagtaaagaaaattaaaaatcttcct

15  gtagcaatgacgattaatcagaaaagattcaatgggaattctcgctccacggtaggaaca
gtttcagatatatatgcttctgttagattactgtggtctagaataggcgaaccgtttgtt
ggttattcagatgcatattccttcaatagtcctaagggcatgtgtaaaacttgtgaggga
ttaggatatattgaagacattaacttagatgaattgctagattgggataagtctttaaat
gaaggtgcaatagactttccttcttttggaccagacaaagagcgtggtaaagcctatcga

20  gataat

Sequence 2660
MDFINITGASQNNLKNIDVNIPKHLVTVFTGRSGSGKSSLVFNTVAAESEQLLNESYSSY
IQFHLNQQPRPKVKKIKNLPVAMTINQKRFNGNSRSTVGTVSDIYASVRLLWSRIGEPFV

25  GYSDAYSFNSPKGMCKTCEGLGYIEDINLDELLDWDKSLNEGAIDFPSFGPDKERGKAYR
DS

Sequence 2661
Contig_0796_pos_728_342,

30  putative peptide of unknown function
atggctattgtaaataaggttattattgttgaaggtaaatcggataagaaaagagtacaa
caagtaatcgctgaacctgcaaatatcatttgtacacatggcactatgagtatagataag
atagacaacatgatagaaacactttatgacaaacaagtttatgttcttgccgattctgat
gatgagggtgaaaaaattagaaaatggtttaaacgttatttaagcgaaagtgaacatatt

35  tatgttgataaaacgttttgtgaggttgctaagtgtcctaaaaattatttagcacatgta
ttaagtagatatggttttaatgtaaaaaaagaaaagaaacttatgaataatttaaaaact
gaaaggctagttttagtaaatgaataa

Sequence 2662

40  MATVNKVTIVEGKSDKKRVQQVIAEPANIICTHGTMSIDKIDNMIETLYDKQVYVLADSD
DEGEKIRKWFKRYLSESEHIYVDKTFCEVAKCPKNYLAHVLSRYGFNVKKEKKLMNNLKT
ERLVLVNE*

Sequence 2663

45  Contig_0798_pos_5428_4553,
putative peptide of unknown function
atgatattgaattcaaaagttaaaggtattattgctatattgatttcagctgtgggtttt
agttttatgtcagtctttttagattggccggtgatttaccagtctttcaaaaatctcta
gctagaaattttgtagccatgtttataccattattttttatttataaatataggcaacct

50  atgtttggaaaattaagtagtcaacccctactcatctcacgttcaacacttgggttaatt
ggtgtcttacttaatatctacgcaattgatcacatggtattaagtgatgctgatacatta
atgaaattaaatccttttggacaattgttcttagtttaattttttttacatgaaaaggta
cgaaaatatcaaatcacggcgatgattattgctataatagggatgctattaattgttaaa
ccagaattttcatcatcagttattccttcaatagcaggattactatccggtattttttgca

55  gcttctgcctacacatgtgttagagcactcagcactcgtgaaaaaccttatacgatagtg
ttttattttttcattattctcagttgtagttcttatacctttttcaatatttacttataca
cctatgacaacaattcaaattctttttcttactcggcgctggattatcagcagctgtagga
caaattggtataacattggcttatagttttgctccagcaaaagatatctccatcttcaca
tatgcgtctataatatttactgcattatttggatttattctgtttggagaatcacctgat

atgtttgcaacagtaggatatattgtcattatcggagcaagttactatatgtttgataaa
gcaagacgtgaaacaactataaatcaaaataattaa

Sequence 2664

5   MILNSKVKGIIAILISAVGFSFMSVFFRLAGDLPVFQKSLARNFVAMFIPLFFIYKYRQP
MFGKLSSQPLLISRSTLGLIGVLLNIYAIDHMVLSDADTLMKLNPFWTIVLSLIFLHEKV
RKYQITAMIIAIIGMLLIVKPEFSSSVIPSIAGLLSGIFAASAYTCVRALSTREKPYTIV
FYFSLFSVVVLIPFSIFTYTPMTTIQILFLLGAGLSAAVGQIGITLAYSFAPAKDISIFT
YASIIFTALFGFILFGESPDMFATVGYIVIIGASYYMFDKARRETTINQNN*

10

Sequence 2665
Contig_0798_pos_3629_2250,
is similar to (with p-value 0.0e+00)
>sp:sp|P94408|YCLF_BACSU HYPOTHETICAL 53.3 KD PROTEIN IN SFP
15  -GERKA INTERGENIC REGION. >gp:gp|Z99106|BSUB0003_15 Bacillus
subtilis complete genome (section 3 of 21): from 402751 to
611850. NID: g2632653. >gp:gp|D50453|D50453_69 Bacillus subt
ilis DNA for 25-36 degree region containing the amyE-srfA re
gion, complete cds. NID: g1805369.
20  atgagagcaatgcttattttttatatgtattttgcactccaggaaaacggacttggaatg
gacaaaacaacagccatgtctatcatgtctgtatatggatcattaatttatatgtcttca
attccaggtggctggatagctgatagaattacaggtacccgaggtgccacattaatcgga
gctatattaattattataggtcacatatgtttaagccttccttttgcaatggtaggttta
ttcacttctatgttctttattattgtaggatcaggtttaatgaaaccaaacatttcaaat
25  attgttggtagactctatccagaaaacgacgtacgcatggatgctggatttgttatcttc
tatatgtcagtgaacatgggtgcactcgtttcaccaattattttacaacactatattgat
attagaaatttccatggcggcttcttgattgcagcaatagggatggctcttggattagtt
tggtacttactatttaaccgaaaaactttgggtagtatcggtatgaaaccgacaaaccca
ttatcttcttctgaaaagaaaaagtacggaacaatcatcggaatcgttgttatagcaata
30  gtattaatccttatgattgcttactttacgcatacgctatcatttaatttaatcagtaat
actgttttaattttaggtattgctttaccaatcatttactttacaacaatgattagaagt
aaagaagctactgatactgaaagatcaagagtaaaagcatttattccgttattca ttta
ggcatgttattttggtcaattcaagacaaggatctaatgtattaaatatctatggtatt
gaaaactctgatatgaaattaaatttatttggttggaaaacacatttggtgaagctatt
35  ttccaaaccattaacccattatttattttattatttgcacccgtggttactcttttatgg
caaaagctaggaaagaaacaacctagcctacctattaagtttgcaattggtactattta
gcaggcgcatcctacatacttatgggagcaatcggtcatatttatggggatacacaattc
tcagttaactgggttattctttcatacgttatctgtgttattggtgagctttgtctctct
ccaactggtagtagtgcagcagttaaattagcacctaaggcatttaacgcacaaatgatg
40  agcctttggttattaactaacgcttcagctcaggccattaacggtacattagttaaatta
attaaaccacttggtcaaaccaattactttatcttcttaggtgttgttgcaaccgtgatt
acgttaattatattagcgtttattcctaagatttctaaagcaatgaaaggtattcgttaa

45  Sequence 2666
MRAMLIFYMYFALQENGLGMDKTTAMSIMSVYGSLIYMSSIPGGWIADRITGTRGATLIG
AILIIIGHICLSLPFAMVGLFTSMFFIIVGSGLMKPNISNIVGRLYPENDVRMDAGFVIF
YMSVNMGALVSPIILQHYIDIRNFHGGFLIAAIGMALGLVWYLLFNRKTLGSIGMKPTNP
LSSSEKKKYGTIIGIVVIAIVLILMIAYFTHTLSFNLISNTVLILGIALPIIYFTTMIRS
50  KEVTDTERSRVKAFIPLFILGMLFWSIQEQGSNVLNIYGIENSDMKLNLFGWKTHFGEAI
FQTINPLFILLFAPVVTLLWQKLGKKQPSLPIKFAIGTILAGASYILMGAIGHIYCOTQF
SVNWVILSYVICVIGELCLSPTGSSAAVKLAPKAFNAQMMSLWLLTNASAQAINGYLVKL
IKPLGQTNYFIFLGVVATVITLIILAFIPKISKAMKGIR*

55  Sequence 2667
Contig_0798_pos_1595_672,
is similar to (with p-value 7.0e-45)
>sp:sp|P39074|BMRU_BACSU BMRU PROTEIN. >gp:gp|L25604|BACBMRU
RBE_1 Bacillus subtilis bmrU, multidrug efflux transporter (

bmr) and its regulator (bmrR) genes, complete cds, and branc
hed-chain 2-oxo acid dehydrogenase (bfmB) gene, 3' end. NID:
g2558636. >gp:gp|D84432|BACJH642_251 Bacillus subtilis DNA,
283 Kb region containing skin element. NID: g2627063. ·gp:g
p|Z99116|BSUB0013_111 Bacillus subtilis complete genome (sec
tion 13 of 21): from 2395261 to 2613730. NID: g2634723.
atgaaacaaccgtataaccatggtgttcttttctatcatgaacatagcggtttgaaagat
atacataatggcataggagaagttgcaaatctcttagttcaatgtgtaaacacctctct
cttcaactcagtgaaaataaaggcgatattattaaatattgtaaatctattaaaaatgaa
aattatagctctgatgtagacgttttatttattttaggtggagatggtacacttaatgaa
ctagtaaatggcgttatgcagtatcagttaaatttaccaatcggtgtaataccaggtggt
acctttaacgattttacaaaaacacttcaactgcaccctaattttaaaacagctagtgag
caattattaacatcacatgctgaatcatatgatgtgttaaaagtgaacgacttatatgta
cttaatttcgttggacttggcttaatagtacaaaatgcagagaatgttcaagatġgttct
aaagatatattcggtaaattcagctatattggatcaaccgttaaaacgttattaaatcct
gttaaatttgatttctcattgactgttgatggtgaaacaaaagaaggcaatacttcgatg
atgttaatagcaaacggtcccaatataggtggtggacaaattccgctaaccgatttatcg
ccacaagatggaagagcaaacacatttgtatttaatgatcaaacactaaatatattgaat
gatatattaaaaaaacgtgatagtatgaattggaacgaaatcacacaaggtattgatcac
atatcaggtaagcacatcacactctcaacaaaccctagtatgaaagtggatattgatggc
gaaattaatttagaaacaccaattgagattcaagtattacccaaagcgatacaacttctt
actgcaactgaacaaaataattaa


Sequence 2668
MKQPYNHGVLFYHEHSGLKDIHNGIGEVAKSLSSMCKHLSLQLSENKGDIIKYCKSIKNE
NYSSDVDVLFILGGDGTLNELVNGVMQYQLNLPIGVIPGGTFNDFTKTLQLHPNFKTASE
QLLTSHAESYDVLKVNDLYVLNFVGLGLIVQNAENVQDGSKDIFGKFSYIGSTVKTLLNP
VKFDFSLTVDGETKEGNTSMMLIANGPNIGGGQIPLTDLSPQDGRANTFVFNDQTLNILN
DILKKRDSMNWNEITQGIDHISGKHITLSTNPSMKVDIDGEINLETPIEIQVLPKAIQLL
TATEQNN*


Sequence 2669
Contig_0798_pos_0_315,
putative peptide of unknown function
atgtttgtgaattatttcacaatatctaaggagtggttgtatatgttatctgtaactaaa
aaaaatacatatgaatcaaacaaagatgaagtcacacaaatgattgattcattagcagaa
aaaggacaagaagctctaaaagaactatctaaaaaatcacaacatgagattaatgacatt
gtacatcagatgagcatggctgctgttgatcagcatatgcatttagctaaactagcttac
gacgaaacaggtagaggtatttatgaagacaaagctatcaaaaatttatatgcctcagag
tacatatggaattcA


Sequence 2670
MFVNYFTJSKEWLYMLSVTKKNTYESNKDEVTQMIDSLAEKGQEALKELSKKSQHEINDI
VHQMSMAAVDQHMHLAKLAYDETGRGIYEDKAIKNLYASEYIWNS


Sequence 2671
Contig_0799_pos_5458_5027,
putative peptide of unknown function
atgttcaaaaatatattattaccctatgatttcgaaaatgattttagtgctatccctgac
tatttagaaaaagtcaccgatgaagattcagttgttgtaatttatcacgttgtaacagaa
aatgatcttgcaattagtgtcaagtattataataagcataaagaagatattattagagaa
aaagagaaaaaactcactccattttttacgtgaattagaaaaaagagatattcaatataaa
atagatgtagattttgggcatattaaagatacaatcttagaaaaaattacttctggagat
ataaataatggtgaatttgatttagtaattatgagtaatcatagagtcgatttgaatatt
aaacatgttttaggagatgttacacataagattgctaaaagaagttctgtcccagtacta
attgttaaataa


Sequence 2672
MFKNILLPYDFENDFSAIPDYLEKVTDEDSVVVIYHVVTENDLAISVKYYNKHKEDIIRE

KEKKLTPFLRELEKRDIQYKIDVDFGHIKDTILEKITSGDINNGEFDLVIMSNHRVDLNI
KHVLGDVTHKIAKRSSVPVLIVK*

Sequence 2673
Contig_0799_pos_2591_1617,
putative peptide of unknown function
atggaacgattttgttgtgtaaatcaaattaactatattcaaatgaatccgttagaagcc
aaatttaaaacgagcgctctaagatcatggaaaactgatcaggcagatgctcataagctt
gcttgtttaggaccgacgctcaaacaaacaggcagcttacctatacatgagttaatattc
tttgaattaagagaacgtgcccgtttcatctagaaatcgagaatgaacaaaatcgactt
aaattttctttatgtttttctattttactgagagaaataaagtggtgacacaagattta
tatgactctatttctagagcatattattgtctcgttgatacacaagctaatcaaaatatg
attgaacactacgcaggattgaacatgaatgatattaatcttttaagggtaacgcctttt
gatgcgaagtcattacctaaccaaagtagtcaattgtatgacacttatattggattatgg
atagatggtttggacgagattgaaatacgagagattgtaaacagcttatttcaatatatt
caacataaagatggctataagttgaaaattttaactaagagtagagataatcttacggaa
aatcttatagatgaagttgctcatctcaatgatttatatcaccaagagaaaaaggaaata
agtgatgtaattgaagacgtgatacagaataaaaaagaaacaatcattgatattgaaaca
gtaccgtttgaagaagatcttgtaagcgttatttcaaaattaagagttgtagtagattta
tctttagagccgaaactttttttacaaatctgttgtattggcgcgggtataccacaaatt
aataaaaagagaacagattatgttaaacatatgcataatggatatattattgatgacata
tcgcaaactgtagaatctttagattattttttggcacatttaaaaaatggaattattctt
atgcatattccatga

Sequence 2674
MERFCCVNQINYIQMNPLEAKFKTSALRSWKTDQADAHKLACLGPTLKQTGSLPIHELIF
FELRERARFHLEIENEQNRLKFSLCFSIFTERNKVVTQDLYDSISRAYYCLVDTQANQNM
IEHYAGLNMNDINLLRVTPFDAKSLPNQSSQLYDTYIGLWIDGLDEIEIREIVNSLFQYI
QHKDGYKLKILTKSRDNLTENLIDEVAHLNDLYHQEKKEISDVIEDVIQNKKETIIDIET
VPFEEDLVSVISKLRVVVDLSLEPKLFLQICCIGAGIPQINKKRTDYVKHMHNGYIIDDI
SQTVESLDYFLAHLKNGIILMHIP*

Sequence 2675
Contig_0799_pos_1108_701,
putative peptide of unknown function
atgttcccaccccgaacacctagtagagatgccactaacccacctcaacaactcttacaa
ttactaggaataccgtcaaatattacccatctcacatacaattttttctgagcatgcatta
ccttggataagtttttatcgtacactatagttttttctatcgctattgcaataatctatatt
tatatcgcaaagaaatatacaaaaatcacactaggttatggtgctttatttggtatagtt
atttggattgtttttcatttaatcttaatgccaattatgcatgtcgtaccgaatgctttt
gatcaaccattttcagaacacctatcagaattttttggacacattgtttggatgatggtt
atagaaatggtcagaaggtatttctataatattcaattaaataaataa

Sequence 2676
MFPPRTPSRDATNPPQQLLQLLGIPSNITHLTYNFSEHALPWISFIVHYSFSIAIAIIYI
YIAKKYTKITLGYGALFGIVIWIVFHLILMPIMHVVPNAFDQPFSEHLSEFFGHIVWMMV
IEMVRRYFYNIQLNK*

Sequence 2677
Contig_0800_pos_2627_2971,
putative peptide of unknown function
gtgacaaaccggaggaaggtgggggatgacgtcaaatcatcatgccccttatgatttgggc
tacacacgtgctacaatggacaatacaaagggcagcgaaaccgcgaggtcaagcaaatcc
catacagttgttctcagttcggattgtagtctgcaactcgactatatgaagctggcatcg
ctagtaatcgtagatcagcatgctacggtgaatacgttcccgggtcttgtacacaccgcc
cgtcacaccacgagagtttgtaacacccgaagccggtggagtaaccatttggagctagcc
gtcgaaggtgggacaaatgattggggtgaagtcgtaacaaggtag

Sequence 2678

VTNRRKVGMTSNHHAPYDLGYTRATMDNTKGSETARSSKSHKVVLSSDCSLQLDYMKLES
LVIVDQHATVNTFPGLVHTARHTTRVCNTRSRWSNHLELAVEGGTNDWGEVVTR*

Sequence 2679
Contig_0802_pos_6032_5214,
is similar to (with p-value 0.0e+00)
>sp:sp|P23966|MENB_BACSU NAPHTHOATE SYNTHASE (EC 4.1.3.36) (
DIHYDROXYNAPHTHOIC ACID SYNTHETASE) (DHNA SYNTHETASE). >pir:
pir|A42715|A42715 dihydroxynapthoic acid synthetase - Bacill
us subtilis >gp:gp|M74521|BACMENAQUI_6 Bacillus subtilis men
aquinone operon, complete cds. NID: g557486. >gp:gp|M74538|B
ACMENAQOP_4 Bacillus subtilis menaquinone operon: menF, menD
, menB and menE genes, complete cds. NID: g1185287.
atgactagacagtgggaaacacttagagaatatgatgaaattaaatatgaatttttcgaa
gggattgccaaagtaacgattaatcgtccagaagtaagaaatgcatttactcctaaaaca
gttgctgaaatgattgatgcattttcacgtgcgcgtgatgatcaaaatgtatcagtaatt
gtattaactggtgaaggggacaaagcgttttgttcaggtggagatcaaaaaaaacgtgga
cacggtggttatgtaggtgaagatgatattcctcgtttaaatgtattagatttacaacgt
ttaattcgtgtgattcctaaaccagtaatagcaatggttagaggctatgcaattggtgga
ggaaatgtacttaatgttgtttgtgatttaactatcgctgcagacaatgctatttttgga
caaactggaccaaaagtaggctcatttgatgctgggtacggttctggctacctagctcgt
atagttggccataaaaaagcaagagaaatctggtacttatgccgtcaatataatgcacag
gaagctttggatatgggcttagtgaatactgtagttccattagaacaagttgaagacgaa
acagttaaatggtgtaaagacatcatgcaacactcaccaactgctttacgtttcttaaaa
gcagcaatgaatgctgatactgatggttttagctggtttacaacaaatggctggagatgcg
ac‌t‌t‌ac‌ttactatactactgatgaagcgaaagaaggacgtgacgcgtttaaag‍aaaa
cgtaatcctgattttgaccaattccctaaattcccataa

Sequence 2680
MTRQWETLREYDEIKYEFFEGIAKVTINRPEVRNAFTPKTVAEMIDAFSRARDDQNVSVI
VLTGEGDKAFCSGGDQKKRGHGGYVGEDDIPRLNVLDLQRLIRVIPKPVIAMVRGYAIGG
GNVLNVVCDLTIAADNAIFGQTGPKVGSFDAGYGSGYLARIVGHKKAREIWYLCRQYNAQ
EALDMGLVNTVVPLEQVEDETVKWCKDIMQHSPTALRFLKAAMNADTDGLAGLQQMAGDA
TLLYYTTDEAKEGRDAFKEKRNPDFDQFPKFP*

Sequence 2681
Contig_0802_pos_1667_123,
is similar to (with p-value 0.0e+00)
>sp:sp|P39634|ROCA_BACSU 1-PYRROLINE-5-CARBOXYLATE DEHYDROGE
NASE (EC 1.5.1.12) (P5C DEHYDROGENASE). >pir:pir|S39731|S397
31 hypothetical protein - Bacillus subtilis >gp:gp|X73124|BS
GENR_77 B.subtilis genomic region (325 to 333). NID: g413923
. >gp:gp|Z99123|BSUB0020_74 Bacillus subtilis complete genom
e (section 20 of 21): from 3798401 to 4010550. NID: g2636240
.
atggtagtacctttcaaaaatgaacctggtattgatttttcagtacagacaaatgttgag
cgttttaatgaagaattaaggaaagtaaaagcgcaactaggacaagatataccacttgtg
attaacggagaaaaacttactaaaactgatacttttaattcagtgaatcctgcgaataca
tcacagctcattgcgaaagtgtctaaagcaacgcaagatgatattgaaaaagctttcgaa
tcagcaaatcatgcgtatcaatcatggaagaagtggtcgcataaggaccgtgcagaatta
ctgttacgtgtagccgcaattatccgtcgtcgaaaagaggaaatttccgctattatggtt
tatgaagccggcaagccttgggatgaagcagttggagatgcagctgagggtattgatttt
atagaatattatgcaagatcaatgatggaacttgcagatggtaagccagtattagacaga
gaaggtgaacataatcgctatttttataaacctattggtacaggcgtgacaattccacca
tggaattttccatttgcaattatggctggtacaaccttagcccctgttgttgcaggtaac
actgtattattaaagcctgctgaggatacagttttgactgcttataaattaatggaaata
ttagaagaagcaggtttaccccaaggtgttgtaaattttgttcctggtgatccaaaagaa
attggagattatttagtcgaccataaagatacacattttgtcacatttacaggatcccga
gctacaggtacacgtatttatgaacgtagtgctgtagtgcaagaaggacaacagttttta

```
aaacgtgttattgcagagatgggtggcaaagatgcgatagttgtagataataatgtagat
acagatttagcggctgaagcaattgttacatctgcttttggtttctctggtcaaaaatgc
tctgcgtgttctcgtgtcatagtccatcaagacgtacatgatgaaatattggaaaaagca
attcaattaactcaaaaattaactttaggtaatactgaagagaacacatttatggggcca
gtaattaatcaaaaacaatttgataaaatcaaaaattatattgaaattggtaaaaaagaa
ggcaaactagagactggtggtggaacagatgattctaccggttatttcattgaaccaacg
attttctccggactacaatctgcggatcgtatcatgcaagaagaatttttggaccagtc
gtaggctttattaaggtcaaggattttgatgaggctattgaagtagctaatgatactgac
tatggtttgacaggcgctgtaattactaatcatcgtgaacattggattaaggctgtgaat
gaatttgatgtaggtaacctttacttgaatagaggttgtacagctgcagtagtgggttat
catccatttggtggattcaagatgtctggtacagatgctaaaacaggaagtccagattac
ttacttaatttcttagaacaaaaagttgtttctgaaatgtttttaa
```

Sequence 2682
```
MVVPFKNEPGIDFSVQTNVERFNEELRKVKAQLGQDIPLVINGEKLTKTDTFNSVNPANT
SQLIAKVSKATQDDIEKAFESANHAYQSWKKWSHKDRAELLLRVAAIIRRRKEEISAIMV
YEAGKPWDEAVGDAAEGIDFIEYYARSMMELADGKPVLDREGEHNRYFYKPIGTGVTIPP
WNFPFAIMAGTTLAPVVAGNTVLLKPAEDTVLTAYKLMEILEEAGLPQGVVNFVPCDPKE
IGDYLVDHKDTHFVTFTGSRATGTRIYERSAVVQEGQQFLKRVIAEMGGKDAIVVDNNVD
TDLAAEAIVTSAFGFSGQKCSACSRVIVHQDVHDEILEKAIQLTQKLTLGNTEENTFMGP
VINQKQFDKIKNYIEIGKKEGKLETGGGTDDSTGYFIEPTIFSGLQSADRIMQEEIFGPV
VGFIKVKDFDEAIEVANDTDYGLTGAVITNHREHWIKAVNEFDVGNLYLNRGCTAAVVGY
HPFGGFKMSGTDAKTGSPDYLLNFLEQKVVSEMF*
```

Sequence 2683
Contig_0804_pos_874_1548,
is similar to (with p-value 3.0e-76)
>sp:sp|P39788|END3_BACSU PROBABLE ENDONUCLEASE III (EC 4.2.9
9.18) (DNA-(APURINIC OR APYRIMIDINIC SITE) LYASE). >gp:gp|L4
7709|BACYPIA_26 Bacillus subtilis (clone YAC15-6B) ypiABF ge
nes, qcrABC genes, ypjABCDEFGHI genes, birA gene, panBCD gen
es, dinG gene, ypmB gene, aspB gene, asnS gene, dnaD gene, n
th gene and ypoC gene, complete cds's. NID: g1146223. >gp:gp
|U11289|BSU11289_3 Bacillus subtilis 168 asparaginyl-tRNA sy
nthetase (asnS) and endonuclease III (jooB) genes, partial c
ds and DnaD protein (dnaD) and (jooC) genes, complete cds.
NID: g533096. >gp:gp|Z99115|BSUB0012_174 Bacillus subtilis c
omplete genome (section 12 of 21): from 2195541 to 2409220.
NID: g2634478.
```
atggagagaatcctaatgataagtaagaaaaaagcattacaaatgattgacgttatagca
gatatgtttcctaatgcagaatgcgaattaaaccatagaaatgcattcgatcttacaata
gctgtattattatcagcacagtgtactgataatctagtcaatcgtgtcactcaatcatta
tttagaaaatatcgaacacctgaagattatttaaatgtgagtgatgaagaattacaaaat
gatatacgctctattggattatatcgcaataaagccaaaaatataaaaaaattatgccac
tctttaattgaacaatttaatggtcaaatcccacaaacacataaagaattagagagtcta
gctggagtggggcgtaaaacagcaaatgttgtaatgagtgtcgcatttggagaaccttct
ttagctgtcgatactcatgttgagagagtttctaaacgtttgggaattaatcgttggaaa
gatagtgtaagacaagtagaagatcgattatgtgatattatcccaagagatagatggaat
aaaagccatcatcaattaatattttttgggagatatcattgtcttgctagaaaacctaaa
tgtgagatatgtccgctgttaaatgattgtagagaaggacaaaaacgacataaagcaaag
ataaaggaggcgtga
```

Sequence 2684
```
MERILMISKKKALQMIDVIADMFPNAECELNHRNAFDLTIAVLLSAQCTDNLVNRVTQSL
FRKYRTPEDYLNVSDEELQNDIRSIGLYRNKAKNIKKLCHSLIEQFNGQIPQTHKELESL
AGVGRKTANVVMSVAFGEPSLAVDTHVERVSKRLGINRWKDSVRQVEDRLCDIIPRDRWN
KSLHQLIFFGRYHCLARKPKCEICPLLNDCREGQKRHKAKIKEA*
```

Sequence 2685

Contig_0804_pos_1553_1885,
putative peptide of unknown function
atgattgaaaaacaggatttcaatcatatagaggaccaacttgatcaactagcaagtaat
aaacaactcaaaacaccagaagctagggaactttttagatagttatttcgatttaattatt
5   aattattttaaacaaataaataacatagatgaaattcattttaatcaactcgatacatat
ccagtagttccaatgaattttgatgaacgctatcattatatggttgcacgtaaacaccat
tttatgggctatcgtcaaatgaaaacattgaaatcagaattaataaaaatgaatgcatct
tatctaattagaaagcaacgtcaacaaaaataa

10   Sequence 2686
MIEKQDFNHIEDQLDQLASNKQLKTPEARELLDSYFDLIINYFKQINNIDEIHFNQLDTY
PVVPMNFDERYHYMVARKHHFMGYRQMKTLKSELIKMNASYLIRKQRQQK*

Sequence 2687
15   Contig_0804_pos_5054_0,
is similar to (with p-value 1.0e-16)
>gp:gp|AF076683|AF076683_2 Staphylococcus aureus oligopeptid
e transporter putative substrate binding domain (opp-1A), ol
igopeptide transporter putative membrane permease domain (op
20   p-1B), oligopeptide transporter putative membrane permease d
omain (opp-1C), oligopeptide transporter putative ATPase dom
ain (opp-1D), and oligopeptide transporter putative ATPase d
omain (opp-1F) genes, complete cds; and unknown gene. NID: g
3800817.
25   atgctcaaacgtacaattaaattcatactttatttaatcgtaagttcgtttattatcttc
attttagttgagaagacatctggtaatccagcgattctgtatctacaacgtcatggttat
acgtcgattacgcaagacaatattgaagcggcacaacatcaacttggcttaggacaacat
gtgttactaagatatatcgattgggttggacatgcactcacgggcaacttaggata:ggc
ttta;ta:;gaacgaagcagttaccgctatgataatggaagccatcgtgccgacgc':tgtg
30   ctaatcattgtctctagttgtatcatgttgccatttggctatattgttggttacticgtt
gggacgcgtccgcatacacgttacgctaatggaattcgtggattcgcccaagtgatgacc
tcaatgccagaatactggttagctatttttattcatttattatttaggcgtacgttggcaa
ttgttaccatttgtaggtagtgattcatggcaacacttgtgctgccaatcttcacaatt
gttgttatagaagggtgtcatatcttattgatgacagcacatctgattacacaaacgtta
35   gatcaagatgcgtatcaactggcgcagttaagacattttttcgttaaaagcgcgtatcatc
gtacaaattaaagagatatttgcaccac

Sequence 2688
MLKRTIKFILYLIVSSFIIFILVEKTSGNPAILYLQRHGYTSITQDNIEAAQHQLGLGQH
40   VLLRYIDWVGHALTGNLGYGFSTNEAVTAMIMEAIVPTLVLIIVSSCIMLPFGYIVGYFV
GTRPHTRYANGIRGFAQVMTSMPEYWLAILFIYYLGVRWQLLPFVGSDSWQHFVLPIFTI
VVIEGCHILLMTAHLITQTLDQDAYQLAQLRHFSLKARIIVQIKEIFAPX

Sequence 2689
45   Contig_0804_pos_3696_3097,
is similar to (with p-value 5.0e-33)
>gp:gp|AF068901|AF068901_4 Streptococcus pneumoniae penicill
in-binding protein 2b (pbp2b), RecM (recM), D-Ala-D-Ala liga
se (ddl), D-Ala-D-Ala adding enzyme (murF), MutT (mutT), cel
50   l division protein FtsA (ftsA), cell division protein FtsZ (
ftsZ), YlmE (ylmE), YlmF (ylmF), YlmG (ylmG), YlmH (ylmH), c
ell division protein DivIVA (divIVA), and isoleucine-tRNA sy
nthetase (ileS) genes, complete cds; and unknown gene. NID:
g4009462.
55   atgataaacgtaacattagagcaaattaaaaactggatagattgtgaaattgatgaaaaa
catttaaaaaaaacaataaatggcgtttcaattgattcacgaaaaatcaatgaaggggcg
ttatttataccttttaaaggtgagaatgttgatggccatcgtttatcacacaagctttg
aacgatggtgctggagctgttttttagtgaaaaagagaataaacattctgaagggaaccaa
ggtcctattatttgggtagaagatactttaatagccttacagcaattggcaaaagcatat

659

```
ctaaatcatgtaaatcctaaggtgatagcggttactggttctaatggaaaaacaa aaca
aaagacatgattgaaagtgtattatcaactgaatttaaagttaagaaaacacaaggaaat
tataataatgaaattggaatgccgttgactttactagaacttgatgaagacacagaaatt
tctattctagaaatggggatgtcaggttttcatcaaatagagttgttatctcatatcgca
caacctgatatagcggtcatcacaaacattggcgaatcacatatgcaagatttaaaataa
```

Sequence 2690
MINVTLEQIKNWIDCEIDEKHLKKTINGVSIDSRKINEGALFIPFKGENVDGHRFITQAL
NDGAGAVFSEKENKHSEGNQGPIIWVEDTLIALQQLAKAYLNHVNPKVIAVTGSNGKTTT
KDMIESVLSTEFKVKKTQGNYNNEIGMPLTLLELDEDTEISILEMGMSGFHQIELLSHIA
QPDIAVITNIGESHMQDLK*

Sequence 2691
Contig_0804_pos_2647_2078,
is similar to (with p-value 3.0e-57)
>gp:gp|Y17795|SAU17795_2 Staphylococcus aureus prfA, pbp2 ge
nes. NID: g3955029.
```
atggtaggttgtcgcattagatatgaatactatcgttcggcttatggacatggtgtgtca
ggcgtaaacatgggtgcgaaaactggtactggtacgtatggacaagaaatatacgaaaag
tataattacctgataatgctgccaaagatgtttggattaatggtttcagtccagaatat
actatgtcggtatggatgggcttcaataaggttaaacaatatggaacaaattcatttatc
ggtcattcagaacaagattatccacaatacttgtatgaagatgtgatgtctagtatctca
tctaaagatggtgaagatttcaaaaagcctaatgatgtacaaggaagttcaccggacagt
ctatctgtatcaggtcattctgataataatactactaaccgtagtgttcatggaagtagc
gatacatcttcttcatcaaatggtggctctaactcagcatcaagtggaaacaactcgaat
agttcgaatggtaccagtcaaggtaactcaggcaatgcatttacacgtctgttcaattta
aactctatattcgattataaagtttcataa
```

Sequence 2692
MVGCRIRYEYYRSAYGHGVSGVNMGAKTGTGTYGQEIYEKYNLPDNAAKDVWINGFSPEY
TMSVWMGFNKVKQYGTNSFIGHSEQDYPQYLYEDVMSSISSKDGEDFKKPNDVQGSSPDS
LSVSGHSDNNTTNRSVHGSSDTSSSSNGGSNSASSGNNSNSSNGTSQGNSGNAFTRLFNL
NSIFDYKVS*

Sequence 2693
Contig_0804_pos_913_473,
putative peptide of unknown function
```
atgcttttttcttacttatcattaggattctctccattcaaccagtcaaatttttcagta
gctaatttttacaaagatgttcacatgactgtttttcaatggatattctctgcattaact
ttatggattgggggttagtttattttttgactttaggattgattatcgctcaactcaacgat
attcaaaaagcaagtagttttgccaatttacttaatattacactagctatattaggaggt
ctatggtttccagtatacacgtttcctgattggcttcagtcgatttctaaacacatgcca
acatataatttaaagctacttgctatagatttagcgcaaaataaagggggtgaatatagaa
gcgtttggctatctcgtggtctattgtataatctttgtgagtattgctttattcatgaat
aagaaaggagatgtacactaa
```

Sequence 2694
MLFSYLSLGFSPFNQSNFSVAHFYKDVHMTVFQWIFSALTLWIGVSLFLTLGLIIAQLND
IQKASSFANLLNITLAILGGLWFPVYTFPDWLQSISKHMPTYNLKLLAIDLAQNKGVNIE
AFGYLVVYCIIFVSIALFMNKKGDVH*

Sequence 2695
Contig_0804_pos_0_452,
putative peptide of unknown function
```
gtgagaataggggaattaagtactttaatatatttaatatttccgattttagctatattt
gttgacaaacgtggtaatttcttaacttatttaattgtttgtactattttttataatcagc
tatgtgacgatgattatattttataaataccttagtgatagtattttatattcattactg
gttattcattatttaggaatctttatttttgtctatagtgtcaatcctatgaatagttttg
```

```
tttttcttctatagtgctttcgccttacctttatttttaatgttcgtgttgtatcaaaa
gaatttataacctttttaatagctatgataagttgtttaatactaacttatatatttaat
ccaacatttgtggttccattaagtgcattttatttggttatattaattgttgctgtaggt
aattttaaaaatagagacgaacgaattatGAC
```

Sequence 2696
VRIGELSTLIYLIFPILAIFVDKRGNFLTYLIVCTIFIISYVTMIIFYKYLSDSILYSLL
VIHYLGIFYFVYSVNPMNSLFFFYSAFALPFIFNVRVVSKEFITFLIAMISCLILTYIFN
PTFVVPLSAFYLVILIVAVGNFKNRDERIMT

Sequence 2697
Contig_0805_pos_575_1141,
putative peptide of unknown function
```
atgaaaggtttaattattataggagtgctcaagtagggtctcatacgaacgctttatca
aaatatttaaaaggtcaactcggcgaacatgatgttgaggtggaaatctttgacctagct
gagaaacccattcatcaattggattttgctggtacaacacaagcagttgatgaaattaaa
aacaatgtcaaatctttacaaaataaagcaatggaagcagatttcttaattttaggaacg
ccaaattatcatggatcgttttcaggtattcttaaaaatgcacttgaccaccttaatatg
gaccatttcaaaatgaaacccgtgggactcatttgcaatagtggaggaatagtaagttct
gagccattatcacacttgagagtcatcgtacgtagtttacttggtattgctgtaccaacg
caaattgctacacatgattctgattatgctaaattagaagatggtaccttatacttagaa
gataatgaatttcaactacgttcaaaattgtttgttgatcaaattgtatccttcgtaaca
aatagtccatatgaacacttaaaataa
```

Sequence 2698
MKGLIIIGSAQVGSHTNALSKYLKGQLGEHDVEVEIFDLAEKPIHQLDFAGTTQAVDEIK
NNVKSLQNKAMEADFLILGTPNYHGSFSGILKNALDHLNMDHFKMKPVGLICNSGGIVSS
EPLSHLRVIVRSLLGIAVPTQIATHDSDYAKLEDGTLYLEDNEFQLRSKLFVDQIVSFVT
NSPYEHLK*

Sequence 2699
Contig_0805_pos_1585_2421,
is similar to (with p-value 1.0e-95)
>gp:gp|Y17116|SEY17116_1 Staphylococcus epidermidis gene enc
oding fibrinogen-binding protein, complete CDS. NID: g320154
9.
```
atgattaataaaaaaaataatttactaactaaaaagaaacctatagcaaataaatccaat
aaatatgcaattagaaaattcacagtaggtacagcatctattgtaataggtgcaacatta
ttgtttggtttaggtcataatgaggccaaagctgaggagaattcagtacaagactttaaa
gattcgaatacggatgatgaattatcagatagtaatgatcagtctagtgatgaagaagag
aatgatgtaattaataataatcagtcaataaactctgatgataataaccaaataaataaa
aaagaagaaacgaataacaacgatggtatagaaaaaagctcagaagatagaacagagtca
acaacaaatgtagatgaaaacgaagcaacatttttacaaaagtcccctcaagataatact
catcttacagaagaagaggtaaaagaaccctcatcagtcgaatcctcaaattcatcaatt
gatactgcccaacaaccatctcacacaacaataaatagagaagaatctgttcaaacaagt
gataatgtagaagattcacacgtatcagattttgctaactctaaaataaaagagagtaac
actgaatctggtaaagaagagaatactatagagcaacctaataaagtaaaagaagattca
acaacaagtcagccgtctggctatacaaatatagatgaaaaaatttcaaatcaagatgag
ttattaaatttaccaataaatgaatataaagtaacgaaacttagcgtcactttcttcatt
gaaaagaaccgtgaaatacttttgactttcatatcaattctccttatgaattattaa
```

Sequence 2700
MINKKNNLLTKKKPIANKSNKYAIRKFTVGTASIVIGATLLFGLGHNEAKAEENSVQDFK
DSNTDDELSDSNDQSSDEEENDVINNNQSINSDDNNQINKKEETNNNDGIEKSSEDRTES
TTNVDENEATFLQKSPQDNTHLTEEEVKEPSSVESSNSSIDTAQQPSHTTINREESVQTS
DNVEDSHVSDFANSKIKESNTESGKEENTIEQPNKVKEDSTTSQPSGYTNIDEKISNQDE
LLNLPINEYKVTKLSVTFFIEKNREILLTFISILLMNY*

Sequence 2701

Contig_0805_pos_4398_4919,
putative peptide of unknown function
atgctcattgatatagttgttcttcttattatttgttactttatagtgatagggttcgt
agaggtatttggttatcgatattgcactttgcttcttcaattgtatctttatata.tgcg
5   tcacaacattatcaatctattgcgcaacgtttagttgtctttgtgccatttccgaaaacg
gtggcgtttgacatggtctatactattccttatgatcatttgcaatacagatttgaaaaa
gtgatagcatttattataatatttggtatgtgtaagcttattttgtatctagttgttgtt
acatttgataatataataacgtataaaaagatacatttagtaagtcggatatcgagtgtc
gttttgagtatcatagcggtttttatatatttacaaattggactttatttattatcgcta
10 tatccgcattcatttatacagtaccaattatctcaatcgctattaagtcgagttgtgatt
gaacaaattccttatttatcacaatttattttaaatttataa

Sequence 2702
MLIDIVVLLIICYFIVIGFRRGIWLSILHFASSIVSLYIASQHYQSIAQRLVVFVPFPKT
15 VAFDMVYTIPYDHLQYRFEKVIAFIIIFGMCKLILYLVVVTFDNIITYKKIHLVSRISSV
VLSIIAVFIYLQIGLYLLSLYPHSFIQYQLSQSLLSRVVIEQIPYLSQFILNL*

Sequence 2703
Contig_0805_pos_4994_5461,
20 putative peptide of unknown function
atgacaaaaaaagatgtaattcaattattagaaaaaatagctatatatatggagctaaaa
ggagaaaatacatttaaagtttcagcgtatagaaaagccgcacaaagtctagaggttgat
gagcgtacattagaagagattgatgatgtaacagaacttaaaggcattggaaaaggcgta
ggagaagttattaatgaatttaaaacacaaggtcaatcatcgacccttcaagcacttcaa
25 gatgaagtacctgaagggttagtgccacttttgaaaatacaaggattaggcataccattg
ataattattatatttttcaacaattggatctatccacatatccatcgctttacaatctgg
acaccagtcggcttcaaacttgataattactggatcagtactaccaatagtagaattaaa
ttcttcggtttgatgaatatttttcatcactatctactccttttgtaa

30 Sequence 2704
MTKKDVIQLLEKIAIYMELKGENTFKVSAYRKAAQSLEVDERTLEEIDDVTELKGIGKGV
GEVINEFKTQGQSSTLQALQDEVPEGLVPLLKIQGLGIPLIIIIFFNNWIYPHIHRFTIW
TPVGFKLDNYWISTTNSRIKFFGLMNIFHHYLLLL*

35 Sequence 2705
Contig_0805_pos_3875_2949,
is similar to (with p-value 2.0e-34)
>sp:sp|O07874|RNH2_STRPN RIBONUCLEASE HII (EC 3.1.26.4) (RNA
SE HII). >gp:gp|U93576|SPU93576_1 Streptococcus pneumoniae r
40 ibonuclease HII (rnhB) gene, complete cds. NID: g2209338.
atgggaaatgtcgtatacaaactcacgtcaaaagaaattcaatcattgatggctcaaact
actttgagacgacgaagttacctcaaggtatgaaagctcgtacgagatatcaaaatact
gttatcat tatctatagttctggcaaagtaatgtttcaaggtaagaatgctgaacaactt
gcgagtcaattgctaccaaataaacaatcaacaactggcaaacatacatcatcaaataca
45 actagtattcaatataatcgttttcattgtattggaagcgatgaagcaggcagtggcgac
tattttggtccattgactgtatgtgcagcttatgtgagccaatcacatatcaaaatctta
aaagaacttggtgtagatgattcaaaaaaactaagcgatactaaaatcgtcgatcttgca
gaacagctcattaccttttatcccgcattctttattaacattagataatgttaagtataac
gaacgacaaagtctaggatggtctcaagttaaatgaaagctgtcttacataatgaagct
50 atcaaaaatgtgcttcaaaaaattgagcaagatcaactggattatattgttattgatcaa
tttgcaaagcgagaagtttatcaacattatgcattatcagcattaccttttcctgacaaa
acaaaatttgaaacaaaaggtgaatctaaatcactagcaatcgcggcagcaagcattatt
tctcgttatgcatttgttaaacacatggaccacatctctaaaaaactccatatggaaata
ccaaaaggagcaagtaacaaagtagatttaattgccgctaaagtcattcaaaaatatgat
55 attcaacaacttgatactatttcaaaaaaacattttaaaaacagagataaagcaattcat
cttatgaatcaaaaatacaataaataa

Sequence 2706
MGNVVYKLTSKEIQSLMAQTTFETTKLPQGMKARTRYQNTVINIYSSGKVMFQGKNAEQL

ASQLLPNKQSTTGKHTSSNTTSIQYNRFHCIGSDEAGSGDYFGPLTVCAAYVSQSHIKIL
KELGVDDSKKLSDTKIVDLAEQLITFIPHSLLTLDNVKYNERQSLGWSQVKMKAVLHNEA
IKNVLQKIEQDQLDYIVIDQFAKREVYQHYALSALPFPDKTKFETKGESKSLAIAAASII
SRYAFVKHMDHISKKLHMEIPKGASNKVDLIAAKVIQKYDIQQLDTISKKHFKNRDKAIH
5   LMNQKYNK*

Sequence 2707
Contig_0806_pos_2408_3289,
is similar to (with p-value 0.0e+00)
10   >sp:sp|P37565|YACC_BACSU HYPOTHETICAL 31.8 KD PROTEIN IN FTS
H-CYSK INTERGENIC REGION. >gp:gp|D26185|BAC180K_134 B. subti
lis DNA, 180 kilobase region of replication origin. NID: g46
7326. >gp:gp|Z99104|BSUB0001_71 Bacillus subtilis complete g
enome (section 1 of 21): from 1 to 213080. NID: g2632267.
15   atgacactgattatatagtgagaggtttagcatacggtgggggaaataagagcatctgct
gcaatcacaacagagtcagtacaagaagcacaaacacgtcattatacatggcctactgct
tctgccgctatgggaagaactatgacagctactgttatgatgggtgcaatgttaaaagga
aaccaaaagttaacagttactgttgatggcaaaggtccaattggcagaattattgctgac
gcagatgctcaaggaaatgttcgtgcatatgtagaccatccacaaacgcattttccactc
20   aacgatcaaggtaaattggatgtacggcgagcagttggtactgatggttccattcaggtt
gttaaagatgttggaatgaaagactactttctggtgcgagtccaatagtatcaggtgag
ctaggagatgatttcacatactactatgccacaagtgaacaaacaccatcatcagtagga
ttgggtgtattagttaatccagacaactcaatcaaagcagcgggaggatttattattcaa
gttatgccaggtgctactgatgaaacggtgactaaattagaagaagccattagtcaaatg
25   caacctgtatcgaaattaattgagcaaggacttacacctgaaggaatattaaatgaaatt
ttgggtgaaggtaatgttcaaattttaaattcaacgtcagcgcaatttgaatgtaattgt
agtcatgagaaatttttaaatgctattaaaggtttaggagaggcagaaattcatagcatg
attaaagaggatcatggagctgaagctgtatgtcacttctgtggtaataaatatcagtat
agtgaaagtgaattagaagatttattagaaacaatgaaatag
30

Sequence 2708
MTHDYIVRGLAYGGEIRAYAAITTESVQEAQTRHYTWPTASAAMGRTMTATVMMGAMLKG
NQKLTVTVDGKGPIGRIIADADAQGNVRAYVDHPQTHFPLNDQGKLDVRRAVGTDGSIQV
VKDVGMKDYFSGASPIVSGELGDDFTYYYATSEQTPSSVGLGVLVNPDNSIKAAGGFIIQ
35   VMPGATDETVTKLEEAISQMQPVSKLIEQGLTPEGILNEILGEGNVQILNSTSAQFECNC
SHEKFLNAIKGLGEAEIHSMIKEDHGAEAVCHFCGNKYQYSESELEDLLETMK*

Sequence 2709
Contig_0806_pos_3617_4558,
40   is similar to (with p-value 0.0e+00)
>sp:sp|P37887|CYSK_BACSU CYSTEINE SYNTHASE (EC 4.2.99.8) (O-
ACETYLSERINE SULFHYDRYLASE) (O-ACETYLSERINE (THIOL)-LYASE) (
CSASE).
gtgttttggatggcacaaaaacctgtagattatgttacacaaattattgggaatacacct
45   gtagtcaaattaagaaacgttgttgatgatgatgcagctgatatttatgttaagttagaa
tatcaaaatccaggtggttcggtaaaagatcgtatcgctttagcgatgattgaaaaagct
gagcgtgaagggaaaattaaacctggtgatacaatcgttgagcctacgagtggtaacact
ggtataggtctagcatttgtatgtgctgccaaggggtacaaagcagtttttacaatgcct
gaaacaatgagccaagagcgccgtaacttattaaaagcttatggtgctgaactagtatta
50   acaccaggatctgaagctatgaaaggtgcaataaaaaaagctaaagaattaaaagaagag
cacggctattttgaaccacaacaattcgaaaatccagcaaatcctgaaattcatgaactt
acaactggaccagaattagttgaacaatttgaaggtcgacaaattgatgcattttttagct
ggtgtaggaactggtggtacgttatctggtgttggtaaagtattgaagaaagaatatcca
aatgtggaaatagtagctattgaacctgaagcttctccagtattaagcggtggtgaacca
55   ggccctcataaattacaaggattgggagcaggtttcgtacctgatactttaaatacagaa
gtttatgacagcatcatcaaagtaggtaatgatactgctatggatatggcacgtcgtgtt
gctagagaagaaggtatattagcaggtatttcatctggtgctgcaatatatgctgctatt
caaaaagcaaaagaattaggtaaaggtaaaacagttgtaacagtattaccaagtaatggg
gaacgttacttatcaacaccattatattcatttgataattaa

Sequence 2710
VFWMAQKPVDYVTQIIGNTPVVKLRNVVDDDAADIYVKLEYQNPGGSVKDRIALAMIEKA
EREGKIKPGDTIVEPTSGNTGIGLAFVCAAKGYKAVFTMPETMSQERRNLLKAYGAELVL

5   TPGSEAMKGAIKKAKELKEEHGYFEPQQFENPANPEIHELTTGPELVEQFEGRQIDAFLA
GVGTGGTLSGVGKVLKKEYPNVEIVAIEPEASPVLSGGEPGPHKLQGLGAGFVPDTLNTE
VYDSIIKVGNDTAMDMARRVAREEGILAGISSGAAIYAAIQKAKELGKGKTVVTVLPSNG
ERYLSTPLYSFDN*

10   Sequence 2711
Contig_0806_pos_5681_6160,
putative peptide of unknown function
gtgaaaacttataacaataatcaatctgtgggtaacttcttaagtgtttcagtacaagat
ggtcaaacagttaaacaaggtgaacgtatcatcaattatgatacaaatgggaataaacgc

15   caacaactattgaacaaagtgaatcaagcacaatctcaagttaatgatgattatcaaaaa
gtaaatcaaagtcctaacaatcatcaattacaagttaaattgactcaagatcaaagtgct
ttaaatgaagctcagcagtcattgtcacaatatgacagacaactcaatgacagcatgaat
gcatcatttgatggtaaaattaacattaaaaatgattcagatgtaggcgaagggcaacct
attttgcaattaatttcttcaaatcctcaaattaacgcaactatcacagagtttgatatt

20   aataaaattaaagaaggcgatgaagtaaatgtcactgtaaatagcaaatgtcccatataa

Sequence 2712
VKTYNNNQSVGNFLSVSVQDGQTVKQGERIINYDTNGNKRQQLLNKVNQAQSQVNDDYQK

25   VNQSPNNHQLQVKLTQDQSALNEAQQSLSQYDRQLNDSMNASFDGKINIKNDSDVGEGQP
ILQLISSNPQINATITEFDINKIKEGDEVNVTVNSKCPI*

Sequence 2713
Contig_0806_pos_6623_6060,

30   putative peptide of unknown function
atgagaaaaggaaatcagaatgaagctttagaagaatttatcggaactttattaaaagat
gagcaatattattatgagttagcattttttagaaagtgaaacacaaaatcttgaaatcata
atggagaagatgattaagcaaggaattacaaaatttcgtattgtacctttactcattttt
agtgcaatgcattatatcagtgatattccacaaatacttaaagagatgaaagctcgatat

35   ccacaaattgatagtaaaatgagtgcgcctcttggtacacatccatatatgaaaacatta
gtagaaaatagaattgctgatgaaaaagtcagtgaaggttcaaccaaagcaactatagta
attgcccatggaaatggaagtggacgttttacgaaagcacatgatgaattaaaagcattt
gttaaaacgcttgatagtcatcatcctgtttatgcaagagctttatatgggacatttgct
atttacagtgacatttacttcatcgccttctttaattttattaatatcaaactctgtgat

40   agttgcgttaatttgaggatttga

Sequence 2714
MRKGNQNEALEEFIGTLLKDEQYYYELAFLESETQNLEIIMEKMIKQGITKFRIVPLLIF
SAMHYISDIPQILKEMKARYPQIDSKMSAPLGTHPYMKTLVENRIADEKVSEGSTKATIV

45   IAHGNGSGRFTKAHDELKAFVKTLDSHHPVYARALYGTFAIYSDIYFIAFFNFINIKLCD
SCVNLRI*

Sequence 2715
Contig_0806_pos_5264_4722,

50   putative peptide of unknown function
atgagtacacaggaattaaaagaaatgatattcaaaactcaagtattccatttgttaat.
cattttaacaaactaagacatcagccaaaatggatacttaaatcaatcatagtgatagta
ttagctattattagcgctttcatcacatataatactagtaatgaactattagataaccaa
tcgatagctaatagccaaatggatgaaaatatgtttcgtatgtctacaactataggcgct

55   ttcataggaacaatctttagtgttgttgtagtattttttaatctttctaattatttctaaa
atatttaaatctgacgccaaagcgagcagtttattctcagcagcactctcttattcaatc
attattttaggatttacaactattatttctttaattcaaatagtttttggactaaagata
acagactataagcttgatagcttaaacattttttcaaaggataataaaacactcatggac
atctgctacttcatgatatcattcatttcaataagttttgagtatatattatcggcatat

taa

Sequence 2716
MSTQELKENDIQNSSIPFVNHFNKLRHQPKWILKSIIVIVLAIISAFITYNTSNELLDNQ
SIANSQMDENMFRMSTTIGAFIGTIFSVVVVFLIFLIISKIFKSDAKASSLFSAALSYSI
IILGFTTIISLIQIVFGLKITDYKLDSLNIFSKDNKTLMDICYFMISFISISFEYILSAY
*

Sequence 2717
Contig_0806_pos_833_267,
is similar to (with p-value 3.0e-54)
>sp:sp|P37476|FTSH_BACSU CELL DIVISION PROTEIN FTSH HOMOLOG
(EC 3.4.24.-). >gp:gp|D26185|BAC180K_132 B. subtilis DNA, 18
0 kilobase region of replication origin. NID: g467326. >gp:g
p|Z99104|BSUB0001_69 Bacillus subtilis complete genome (sect
ion 1 of 21): from 1 to 213080. NID: g2632267.
atgatacatggggcatttttcttagcattttcaaataaatcacgaacacgactcgcacca
acaccaacaaacatctcaacgaagtcagatccactaattgagaagaatggtgcaccagct
tcacctgcaaccgcacgtgctaataatgttttacctgtacctggaggcccaacaagtaag
acaccttttggaattcttgaacccatttgtttaaatttcttgttatcttttaagaaatct
acaatttctattaattcttgtttctcttcgtcagctcctgctacatctgagaaacgaact
cgacgtttattactgtcgtacatcttagctttggattttccaaagttcatcatacgacca
ccaccgccgccaccttgggcttggctaaggaagaaaataaataataatgcaatgattaat
acaggaatcagtgtcgttaaaatactaacgaatacactttgttttttcttcttctttaact
gtaaatttaagattatcttgtttcttagctgtatctgtgattttttgtaaatctttttca
ttgttatataaaattgtcgatgagtaa

Sequence 2718
MIHGAFFLAFSNKSRTRLAPTPTNISTKSDPLIEKNGAPASPATARANNVLPVPGGPTSK
TPFGILEPICLNFLLSFKKSTISINSCFSSSAPATSEKRTRRLLLSYILALDFPKFIIRP
PPPPPWAWLRKKINNNAMINTGISVVKILTNTLCFSSSLTVNLRLSCFLAVSVIFCKSFS
LLYKIVDE*

Sequence 2719
Contig_0808_pos_5584_4772,
is similar to (with p-value 1.0e-62)
>gp:gp|AJ223960|LLCAJ3960_4 Lactococcus lactis cremoris MG13
63-inv3 chromosomal inversion junction DNA. NID: g3647234.
atgaaagatcaattgagaggtaatgtcgtgattatacatcaacctgcagaagaaatgcca
ccgggcggtgcgaaaggtatgatagatgacggcgtattagaaggggtggaccatgtttta
ggtgcacatgtgatgagtatgatggaaacaggtaaaatatactatcgtgaaggttttgtt
caaacgggacgcgcttattttagacttgttgtgaaaggtcagggtggtcacggctcatcc
ccacatacatcgaatgacgctattgtagcaggtgcgcattttgtaacgaccgcacagacc
attgtttcaagacgcttaaatccgtttgaaacgggtgtagtcacaataggttccttcgac
gggaaaggacaattcaatgtgattaaagatacaataacaatcgaaggtgatgtacgtgca
ctgactgatgatacaagagacaatattcagaatgaaatgacacgactagtcagaggatta
gaagagatgttcggagtgatttgtgattttgaatttaaaaaggattatccagctctttac
aatgatcctgaatttacttcttatgtagcaacaactttaaaaaatgctaagctggatgac
ataaaagcaatagatatttgtgagccgcagccaccctcagaagatttcgcgtttttatgcg
ttagaaagaccttcaacatttatttattcgggtgcagcaccagaagatggacctatgtac
cctcaccatcatcctaagtttaatatcaatgaaacatctatgcttgtagttgcagaggca
gtgggtacaattgtattagattatttgaaataa

Sequence 2720
MKDQLRGNVVIIHQPAEEMPPGGAKGMIDDGVLEGVDHVLGAHVMSMMETGKIYYREGFV
QTGRAYFRLVVKGQGGHGSSPHTSNDAIVAGAHFVTTAQTIVSRRLNPFETGVVTIGSFD
GKGQFNVIKDTITIEGDVRALTDDTRDNIQNEMTRLVRGLEEMFGVICDFEFKKDYPALY
NDPEITSYVATTLKNAKLDDIKAIDICEPQPPSEDFAFYALERPSTFIYSGAAPEDGPMY
PHHHPKFNINETSMLVVAEAVGTIVLDYLK*

Sequence 2721
Contig_0808_pos_4094_3633,
is similar to (with p-value 3.0e-18)
>sp:sp|P38049|YIXC_BACSU HYPOTHETICAL 18.8 KD PROTEIN IN PBP
F 5'REGION. >pir:pir|B40614|B40614 hypothetical protein X (p
bpF 5' region) - Bacillus subtilis >gp:gp|L10630|BACPBPF_2 B
acillus subtilis penicillin-binding protein (pbpF) gene, 5'
end. NID: g304158. >gp:gp|Z99109|BSUB0006_86 Bacillus subtil
is complete genome (section 6 of 21): from 999501 to 1209940
. NID: g2633260. >gp:gp|Y14083|BSY14083_4 Bacillus subtilis
chromosomal DNA, region 76-78 degrees: between glyB-aprE. NI
D: g2226224.
atgaagggatgtcttatgttaattagcgatacaaataaaacatatataatagaagaaact
agtaattcattcacaattgaaaagaataatgatcagcagcactacgaagtattagaatca
atcaacagcttatctaatgattcattttgtgtgttaaatcacttattcgtcaatggaggt
aatgaagaggtttttgagtcacggttttttaaagcgaaatcaacatttgcaagatgtgcct
ggttttaaagcgttaagatttcttagaccggtagtcaaagggagacattacattatcatc
acgctatggaacagtagacaagctttctatgattggcaaaattcacaagcatatgcgcaa
actcataaaaaacgtggaactcaaaaaggtgttgatcatcgtatagtcaatagagattta
tcctataatataagaatagagttagaaagtcttaataactaa

Sequence 2722
MKGCLMLISDTNKTYIIEETSNSFTIEKNNDQQHYEVLESINSLSNDSFCVLNHLFVNGG
NEEVFESRFLKRNQHLQDVPGFKALRFLRPVVKGRHYIIITLWNSRQAFYDWQNSQAYAQ
THKKRGTQKGVDHRIVNRDLSYNIRIELESLNN*

Sequence 2723
Contig_0808_pos_1986_1399,
is similar to (with p-value 4.0e-51)
>sp:sp|P37470|SP5C_BACSU PROBABLE PEPTIDYL-TRNA HYDROLASE (E
C 3.1.1.29) (PTH) (STAGE V SPORULATION PROTEIN C). >gp:gp|D2
6185|BAC180K_116 B. subtilis DNA, 180 kilobase region of rep
lication origin. NID: g467326. >gp:gp|Z99104|BSUB0001_53 Bac
illus subtilis complete genome (section 1 of 21): from 1 to
213080. NID: g2632267.
gtggaggtaacaataatgaaatgcattgtcggtcttggcaacattggtaaacgttttgaa
ttaacaagacataatattggtttcgaagttgtcgatgatattctagaacgccaccaattt
actttagacaaacaaaaatttaaaggtgcatatactattgaacgtttaaacggcgaaaaa
gtattatttattgagccaatgaccatgatgaacttatctggtcaagctgtagcccccttta
atggattattataatgtcgatgttgaagatttgatcgtttatatgacgatttagattta
gaacaaggacaagtgcgtctgcgccaaaaggggagtgcaggcggtcataatggtatgaaa
tcgataattaaaatgcttggtacagatcaatttaaacgtattcgaattggtgttggccgt
ccaacaaatgggatgtctgttccggactatgttttacaaaaattttcaaaagaagaaatg
atcattatggaaaaggtaattgaacattctgcaagagctgtagaatcttttattgaaagt
tctcgttttgatcatgttatgaatgaatttaatggtgaagtcaagtga

Sequence 2724
VEVTIMKCIVGLGNIGKRFELTRHNIGFEVVDDILERHQFTLDKQKFKGAYTIERLNGEK
VLFIEPMTMMNLSGQAVAPLMDYYNVDVEDLIVLYDDLDLEQGQVRLRQKGSAGGHNGMK
SIIKMLGTDQFKRIRIGVGRPTNGMSVPDYVLQKFSKEEMIIMEKVIEHSARAVESFIES
SRFDHVMNEFNGEVK*

Sequence 2725
Contig_0808_pos_0_1393,
is similar to (with p-value 4.0e-65)
>gp:gp|AF054624|AF054624_1 Lactobacillus sakei transcription
-repair coupling factor (mfd) gene, partial cds; L-lactate d
ehydrogenase (ldhL) gene, complete cds; and unknown genes. N

```
ID.  35 1014.
atgattgcaaattacattagcgaagataatcgttttcaagaattagatgaagtcttcggc
caagagaatattttagttacgggattatccccgtcagcgaaggcaacaattattgctgaa
aaatatttaaaagatcataaacaaatgctactcgtaactaataatttataccaagcagat
```
5
```
aaaatcgaaactgacattttacaatatgtagatgactcagaagtttataaatatcctgtt
caagatataatgactgaagagttttctactcaaagtccacaattgatgagtgagcgtgtt
agaacgttgactgcgttagcccaaggcgaaaaagggttatttattgtgcctttaaacggc
tttaaaaaatggctaacaccggttgatttatggaaagatcatcaaatgacgcttaaagta
ggtcaggatattgatgttgatgcattcttaaataaattagttaatatgggttatcgccga
```
10
```
gaaagtgtagtgtcacatattggtgagttttcattgcgtggtggaatcatagatatatat
ccgttgataggaacacctgtgagaatagagctatttgatactgaagttgattccatcaga
gactttgatgtagaaacacaacgttctaacgataatataaatcaagttgaaatcacaaca
gctagtgactacattattactgatgaagtgatacaacacttacaaaatgaacttaaaaaa
gcatatgaatatacacgccctaaaattgaaaagtccgtacgtaatgatttaaaagagaca
```
15
```
tatgaaagttttaagttgtttgaatctacctttttcgatcatcaattgttacgacgtctt
gtttcatttatgtatgaaaaaccatcaacccttattgactattttcaaaaaaacgcgatt
attgtagttgatgagtttaatcgtattaaggaaacagaagaaacacttacaacagaagtt
gaagatttatgagtaacttaattgagagtggaaatggatttatcggacaaggatttatg
aagtatgaaagttttgacgcattattagagcaacatgccgttgcatatttcacattattt
```
20
```
actcttcgatgcaagtaccattacaacatattattaagttctcttgtaaaccagttcaa
caattttatggtcaatatgacattatgcgctcggaatttcaaagatacgtgcatcaagat
tacactgtcgtagttcttgttgaaactgaaacaaaagttgaacgtattcaatcaatgctt
aatgaaatgcatattccaacagtatcaaatattcacgaagatattgatggtggtcaagtt
gtagtgacggaaggtagtctttctgaaggctttgaattaccttatatgcagttggtagtc
```
25
```
atcacagaaagag
```

Sequence 2726
MIANYISEDNRFQELDEVFGQENILVTGLSPSAKATIIAEKYLKDHKQMLLVTNNLYQAD
KIETDILQYVDDSEVYKYPVQDIMTEEFSTQSPQLMSERVRTLTALAQGEKGLFIVPLNG
30
FKKWLTPVDLWKDHQMTLKVGQDIDVDAFLNKLVNMGYRRESVVSHIGEFSLRGGIIDIY
PLIGTPVRIELFDTEVDSIRDFDVETQRSNDNINQVEITTASDYIITDEVIQHLQNELKK
AYEYTRPKIEKSVRNDLKETYESFKLFESTFFDHQLLRRLVSFMYEKPSTLIDYFQKNAI
IVVDEFNRIKETEETLTTEVEDFMSNLIESGNGFIGQGFMKYESFDALLEQHAVAYFTLF
TSSMQVPLQHIIKFSCKPVQQFYGQYDIMRSEFQRYVHQDYTVVVLVETETKVERIQSML
35
NEMHIPTVSNIHEDIDGGQVVVTEGSLSEGFELPYMQLVVITERX

Sequence 2727
Contig_0810_pos_182_592,
putative peptide of unknown function
40
atgaaacgaatgcgcaaaactgaaagcttggttatcaataatcctgtgttactaat
gaacatgaagacgaagcagatatactgtatataggatttatatctactaaaggtgctatt
ggagaaggtgcagaaagactagaacgacatggtgtaaaagtgaatacgatgcatattcga
caattacatcctttccctaaagatattgttcaacaagctattaataaagcttcgaaagta
atagttgcagaacataattatcaaggacaattatcaagtattttaaaaatgaacacacaa
45
gttaatgataaattagttaatcaaacaaaatacgatgggaaacctttcttaccttatgaa
attgaagaaaaggtttggaaattgctaaagagttaaaggagttggtgtaa

Sequence 2728
MEKRMRKTESLVINNPVLLNEHEDEADILYIGFISTKGAIGEGAERLERHGVKVNTMHIR
50
QLHPFPKDIVQQAINKASKVIVAEHNYQGQLSSILKMNTQVNDKLVNQTKYDGKPFLPYE
IEEKGLEIAKELKELV*

Sequence 2729
Contig_0810_pos_2045_3106,
55
putative peptide of unknown function
atggacaaatttaaatctatgacagaattaaaagaattgactaaagaaggaaaagattgg
gaaatagagtgtgaaaatcgttctagcatagtcactatattagcattacatggcggtgga
attgaacctgccacaactgaattagcctatacaattgcacattgtggcgactataactat
ttttcctttaaaggtatgagaagtaagggaataatgagttacatgtgacttccaccat
```

tatgatgaccaaattgcattagatttagtgagaggtagccaaagaactgtagccatccat
ggttgtgaaggtaatgaaagtgtggcttatataggaggtagtgatgacagactaattgag
ttaatcaccgaatctcttgaagatataggaattagcgtgcgagaagcaccacatcatatt
tctggaactcaagaaaatactgaaaatggagcagtcatctataaaggtaatcaactttac
5   aattatcgaagttttgatcagtatattttcaaaaagttgtaaattatttaaatttgaat
caaaagataaacaatttgattatttgtggtgtaaaaagtgcatatattttaaaagaaaca
agcgaagcatttaagcaagatgcacgtacatattatcaccaactaatagaggttgactcc
ttacaaacattacctgatgatgattatgtgaaaattgctttcaatataaatcgtcagact
catccagacttagatgagaaattagctcttaagtttaaagacgatattaaactagtatca
10   agtgggagagatagtatagatgttattatgccaaatatgactaagggtcaagctttgtct
agattattaaaagaatggcaaatgcctgcttcacatttaatggcatttggagatgcaaat
aacgataaagatatgttggagcttgccgaacatagttatgttatggctaatagtgaagat
caatcattatttaatatagcgagtcatgtggcaccttccaatgatgaacaaggcgtacta
tcaacaatcgaaaatgttgttctcggttattccaataaataa

15

Sequence 2730
MDKFKSMTELKELTKEGKDWEIECENRSSIVTILALHGGGIEPATTELAYTIAHCGDYNY
FSFKGMRSKGNNELHVTSTHYDDQIALDLVRGSQRTVAIHGCEGNESVAYIGGSDDRLIE
LITESLEDIGISVREAPHHISGTQENTENGAVIYKGNQLYNYRSFDQYIFQKVVNYLNLN
20   QKINNLIICGVKSAYILKETSEAFKQDARTYYHQLIEVDSLQTLPDDDYVKIAFNINRQT
HPDLDEKLALKFKDDIKLVSSGRDSIDVIMPNMTKGQALSRLLKEWQMPASHLMAFGDAN
NDKDMLELAEHSYVMANSEDQSLFNIASHVAPSNDEQGVLSTIENVVLGYSNK*

Sequence 2731
25   Contig_0810_pos_4526_5194,
     is similar to (with p-value 2.0e-20)
     >gp:gp|AF012552|AF012552_2 Helicobacter pylori prolipoprotei
     n diacylglycerol transferase (lgt) and NADPH-linked flavin n
     itroreductase (rdxA) genes, complete cds. NID: g2564440.
30   atgattatgaatcagatgaatcaaacgattattgatgcattccattttagacatgcgaca
aaagaatttgaccctacgaaaaaaattagtgatgaagattttaatacgattttagaaaca
ggtagattatctccaagttcactaggtttagaaccttggcactttgtagtggttcaaaat
aaagaattgagagaaaaattgaaagcctatagttggggagcacaaaagcaacttgataca
gcaagtcactttgtattaattttttgctcgtaagaatgtgacggctcatacagattacgtg
35   caacatttacttcgtggcgtcaaaaaatatgaagaaagtacaattccagcagttgaaaat
aaatttgatgatttccaagaaagtttccatattgccgataatgaacgaacattatatgac
tgggcgagtaaacaaacatatattgcattagcaaacatgatgacaagtgctgcattacta
ggtatcgactcatgtccaattgaaggatttgatttagataaagtgactgaaattctttca
gatgagggtgttttagatacggaacaatttggtatttcagttatggtaggctttggttac
40   agagcacaagaacctaaacatggcaaagttagacaaaacgaagacgacatcattagttgg
attgaataa

Sequence 2732
MIMNQMNQTIIDAFHFRHATKEFDPTKKISDEDFNTILETGRLSPSSLGLEPWHFVVVQN
45   KELREKLKAYSWGAQKQLDTASHFVLIFARKNVTAHTDYVQHLLRGVKKYEESTIPAVEN
KFDDFQESFHIADNERTLYDWASKQTYIALANMMTSAALLGIDSCPIEGFDLDKVTEILS
DEGVLDTEQFGISVMVGFGYRAQEPKHGKVRQNEDDIISWIE*

Sequence 2733
50   Contig_0810_pos_4292_3300,
     is similar to (with p-value 0.0e+00)
     >gp:gp|U31175|SAU31175_1 Staphylococcus aureus D-specific D-
     2-hydroxyacid dehydrogenase (ddh) gene, complete cds. NID: g
     1644432.
55   atgacaaaaattatgttttttcggcacaagagcatatgagaaggacatggcattacgttgg
ggaaagaaaaataatatcgatgtcactacatcaacagaacttttaagtgtagatactgtc
gatcaattaaaagattatgacggtgttacaacaatgcagttcggtaaattagaacctgaa
gtttaccctaaattagagtcctatggtattaaacaaattgcacaacgtacggctggattt
gatatgtatgacttagaacttgcaaaaaaacatgaaattattatctcgaatatacctagt

```
tattcacctgaaacaattgctgaatattcggtatctatcgctctgcaactcgtacgaaaa
ttcccaacaattgaaaaacgtgtgcaagcacataatttcacatgggcgtcccctattatg
tctcgtccagtaaaaaatatgactgtagcaatcatcggtacagggcgtattggtgctgca
actggtaaaatctatgctggttttggtgcgagagtagttggttatgatgcatatctaat
```

```
cattctttatctttcttagaatataaagaaacagtagaggatgcaattaaagatgctgat
attatctcattacatgtacccgctaataaagatagtttccatttatttgataacaatatg
tttaaaaatgttaaaaaaggtgccgtttttagtcaatgccgcaagaggagctgtgataaac
acgcctgatttaattgaagcagtaaataatggtacattatcaggtgctgccattgacaca
tatgaaaatgaagctaattatttcacatttgattgttcaaatcaaacgattgacgaccca
atattattagacctaattagaaatgaaaatatttttagttacacctcatattgccttttc
tccgatgaagcagtacaaaatttagtagagggtggtttgaatgcagcattatcagtaatt
aatactggcacatgtgatacgcgattaaactaa
```

Sequence 2734
MTKIMFFGTRAYEKDMALRWGKKNNIDVTTSTELLSVDTVDQLKDYDGVTTMQFGKLEPE
VYPKLESYGIKQIAQRTAGFDMYDLELAKKHEIIISNIPSYSPETIAEYSVSIALQLVRK
FPTIEKRVQAHNFTWASPIMSRPVKNMTVAIIGTGRIGAATGKIYAGFGARVVGYDAYPN
HSLSFLEYKETVEDAIKDADIISLHVPANKDSFHLFDNNMFKNVKKGAVLVNAARGAVIN
TPDLIEAVNNGTLSGAAIDTYENEANYFTFDCSNQTIDDPILLDLIRNENILVTPHIAFF
SDEAVQNLVEGGLNAALSVINTGTCDTRLN*

Sequence 2735
Contig_0812_pos_5666_6487,
is similar to (with p-value 1.0e-57)
>gp:gp|U75480|SMU75480_1 Streptococcus mutans putative HPr(s
  er) kinase (ptsK) and putative prolipoprotein diacylglycerol
   transferase (lgt) genes, complete cds. NID: g3924622.
```
atggctggatattttttcacattatgcttcagaccgtattcaattattagggacaacggag
ttatcattttataatttacttccagatgaagagaagaaaggaagaatgagaaaattatgc
cgacctgaaactccagcgattattgttacacgtgggttagaaccacccgaagaacttata
caagcatctcaagaaacgcatacaccaattattgttgcgaaagatgccacaacgagttta
atgagtaggttaacgacatttctcgaacatgaactcgcgaaaactacttctttgcacggt
gtacttgttgatgtttacggtgtaggtgtactaattacaggagattctggcattgggaaa
agtgaaactgcattagaattagtcaaacgaggccatagattagtggctgatgataatgta
gaaatcaaagaaattactaaggatgaacttgtagggaaaccgcctaaacttatcgaacat
ttgctagagattcgtggtctcggaatcattaatgttatgactttgtttggagcaggatca
atattaactgaaaaacaagttcgattaaatattaatttagaaaattggaataagaataaa
ttatacgatcgtgtaggtcttaatgaagaaacattgaaaattcttgatacggaaatcact
aaaaaaacgataccagttagaccagggcgtaatgtagcagtaattattgaagtagctgct
atgaattatcgtcttaatatcatgggtattaatacagcagttgaatttaatgagagactt
aatgaagaaatcgttcgaaatagtcataaaagtgaggagtaa
```

Sequence 2736
MAGYFSHYASDRIQLLGTTELSFYNLLPDEEKKGRMRKLCRPETPAIIVTRGLEPPEELI
QASQETHTPIIVAKDATTSLMSRLTTFLEHELAKTTSLHGVLVDVYGVGVLITGDSGIGK
SETALELVKRGHRLVADDNVEIKEITKDELVGKPPKLIEHLLEIRGLGIINVMTLFGAGS
ILTEKQVRLNINLENWNKNKLYDRVGLNEETLKILDTEITKKTIPVRPGRNVAVIIEVAA
MNYRLNIMGINTAVEFNERLNEEIVRNSHKSEE*

Sequence 2737
Contig_0812_pos_6493_0,
is similar to (with p-value 8.0e-40)
>sp:sp|P52282|LGT_STAAU PROLIPOPROTEIN DIACYLGLYCERYL TRANSF
ERASE (EC 2.4.99.-). >gp:gp|U35773|SAU35773_1 Staphylococcus
  aureus prolipoprotein diacylglyceryl transferase (lgt) gene
  , complete cds. NID: g1016769.
```
atgaatataacattaggatatatcgatcctgttgcctttagcttgggaccaatccaagtt
cgatggtatggaattattattgcttgtggtatcttacttggatacttcattgcacaagca
gcattgaaacaggttggattacataaagacaccttaatcgatattatattttatagcgcg
```
```

attgttggattcatagttgcgagaatatactttgttacatttcaatggccatattacatg
aatcacttgagtgagataccaaaaatttggcatggtggtattgccatacatggtggctta
attggtggacttatctctgggattattgtttgtaaaatcaaaaatctacatccgtttcaa
ataggagatattgtggc

5

Sequence 2738
MNITLGYIDPVAFSLGPIQVRWYGIIIACGILLGYFIAQAALKQVGLHKDTLIDIIFYSA
IVGFIVARIYFVTFQWPYYMNHLSEIPKIWHGGIAIHGGLIGGLISGIIVCKIKNLHPFQ
IGDIVA

10

Sequence 2739
Contig_0812_pos_6327_5983,
is similar to (with p-value 3.0e-31)
>gp:gp|U75480|SMU75480_1 Streptococcus mutans putative HPr(s
er, kinase (ptsK) and putative prolipoprotein diacylglycerol
 transferase (lgt) genes, complete cds. NID: g3924622.
gtgatttccgtatcaagaattttcaatgtttcttcattaagacctacacgatcgtataat
ttattcttattccaattttctaaattaatatttaatcgaacttgttttttcagttaatatt
gatcctgctccaaacaaagtcataacattaatgattccgagaccacgaatctctagcaaa
tgttcgataagtttaggcggtttccctacaagttcatccttagtaatttctttgatttct
acattatcatcagccactaatctatggcctcgtttgactaattctaatgcagtttcactt
ttcccaatgccagaatctcctgtaattagtacacctacaccgtaa

20

Sequence 2740
VISVSRIFNVSSLRPTRSYNLFLFQFSKLIFNRTCFSVNIDPAPNKVITLMIPRPRISSK
CSISLGGFPTSSSLVISLISTLSSATNLWPRLTNSNAVSLFPMPESPVISTPTP*

25

Sequence 2741
Contig_0813_pos_1018_2097,
putative peptide of unknown function
atgcttgaaaaaacattcgaagtcacgtatacaaatgaacaaaaaattgaattagaagca
caattgttttcaacacaacttttatttcaatttctcttttcgcaaggtaggttagaagaa
gcccgaacatatattttgaatcaatcttacgagatacaacagcataggtgattaggaat
ttacttgcaatgtgttatttgtatctaggtgagtatgatagcgccaaagcaatgtttgaa
gaacttttaaaggaagataattcagacgtgcatgcactttgtcactacacattattactt
tataataaaaaagaaacagaaaaatatcaaaaatatcttaaaatacttaataaagtagta
ccactaaatgacgacgaaacctttaaattaggaatcgtattgagttatttaaaacagtat
cgtgcttctcaaaatttactttatccactttataaaaaaggtaaatttgtctctattcaa
atgtataatgcattgagtttcaattttttattacctaggaaataaagacgaaagtattgag
atgtggaacaagctcactcaaatttctgaagttgatgttggttatgcaccttgggtaatt
gaggaaagtaaaacggtatttgaatcacgagtgttaccattattactagatgataataat
cattatcgactttacggtatttttttacttcatcaattaaatggaaaagaaatactaatg
actgaagatatttggtcaattcttgaatcaatgaatgactatgagaaactttatctcaca
tatttggtacaaggactcacactcaataaattagattttatacacagaggtatgcaaagg
ttgtataattttaagaaattcaaatataacacgtctttatttacagattggattaatcaa
gcagaaatgattatagctgaaaatgtagatttagtagatgtcgatagatatgtagctgca
tttgtttacctatcgtatcgtcgttctagccaaccacttaccaagaggcaattgatggac
gattttaatgtttctagatacaaactgaataaagcaattgaatttatattgagcatataa

50

Sequence 2742
MLEKTFEVTYTNEQKIELEAQLFSTQLLFQFLFSQGRLEEARTYILNQSYEIQQHRVIRN
LLAMCYLYLGEYDSAKAMFEELLKEDNSDVHALCHYTLLLYNKKETEKYQKYLKILNKVV
PLHDPETFKLGIVLSYLKQYRASQNLLYPLYKKGKFVSIQMYNALSFNFYYLGNKDESIE
MWNKLTQTSEVDVGYAPWVIEESKTVFESRVLPLLLDDNNHYRLYGIFLLHQLNGHEILM
TEDIWSILESMNDYEKLYLTYLVQGLTLNKLDFIHRGMQRLYNFKKFKYNTSLFTDWINQ
AEMIIAENVDLVDVDRYVAAFVYLSYRRSSQPLTKRQLMDDFNVSRYKLNKAIEFILSI*

Sequence 2743
Contig_0813_pos_2162_2497,
is similar to (with p-value 6.0e-50)
>gp:gp|AJ223781|SAAJ3781_1 Staphylococcus aureus trxB gene.
5   NID: g3582102.
atgactgaagtagattttgatgtagcaataatcggtgcaggtcctgccggtatgacagca
gcagtatatgcatctcgtgccaatttaaaaactgtcatgattgaacgcggtatgc.aggc
ggtcaaatggcaaacactgaagaagtagagaattttccaggatttgagatgatcacaggt
cctgacttatctactaaaatgtttgaacatgctaaaaaatttggtgcggaataccaatat
10  ggcgatattaaatctgttgaagataaaggcgactataaagttatcaatttagggaatata
gttgttgatgtaaatgtgtgtcacagtattgtttaa


Sequence 2744
MTEVDFDVAIIGAGPAGMTAAVYASRANLKTVMIERGMPGGQMANTEEVENFPGFEMITG
15  PDLSTKMFEHAKKFGAEYQYGDIKSVEDKGDYKVINLGNIVVDVNVCHSIV*


Sequence 2745
Contig_0813_pos_5893_6465,
putative peptide of unknown function
20  atgaaaaaagtattagcactattatttgcgtcaacactcattttaggagcatgtggggac
aaaaatgacgaatctaaaaatgattcttcatctaattctacagatagcgtttctgtagac
aaaaatgacaatgaagataaaacaacagcgaaattcaaaaatgataaattaacaacagac
aattttgatattgaaattttagaagccaaaactgtcaaagcatctgagtacgatgatgac
aagaaaccaagtatcgctatcatctatggtgttaaaaataaaaaagacaaagatttaaca
25  gcatcttcagcatttatcgaatcatttgatatttatcaaaattctaaagatgttaagaga
agattagaaattggcggtggatatgatactgatttaaaggaaaaatatgaagaagatatg
gataacaaaattaacaaagatggaaaagtaaaaggtgttatgttctttaaattgaaagac
actaaaacaccagtaacacttgaagctaaagatccaaatcatagcaataacgaaaaagtg
ggtactaaagaatttaaaattaaagaaaaataa

30
Sequence 2746
MKKVLALLFASTLILGACGDKNDESKNDSSSNSTDSVSVDKNDNEDKTTAKFKNDKLTTD
NFDIEILEAKTVKASEYDDDKKPSIAIIYGVKNKKDKDLTASSAFIESFDIYQNSKDVKR
RLEIGGGYDTDLKEKYEEDMDNKINKDGKVKGVMFFKLKDTKTPVTLEAKDPNHSNNEKV
35  GTKEFKIKEK*


Sequence 2747
Contig_0813_pos_6510_0,
putative peptide of unknown function
40  atgtctcaaaaaattaaagttatagtgcctatcgtacttggattaatcattcttttgggg
attgcttggggtgtatatgcctttgtaacaaatacacctaaaaatgcatatttactgagc
gaaaaaaaagaccgctataaatgtaaaatcttatgttgatcatcgttttagtaacgaaaag
aaattccaaaaaaaattaaaagataattcatatgttaatacgtataatctacatgctaat
gcatctaaggaatatctaaaagatcttggtttacctaaaactattttagatagttctaaa
45  ataactggaactatcggtcatgatccaaaatcaaataaaggaatcatgagcgtatcacct
aaaatattagataaagatattggtaagttccaatggacagcaaatgattcaactcaattc
ttcgaatcacccttattcaagaaaaagtatagcgtcaaaaactcagaactattagaaaca
gctgctcaaatctttgatgaagatccttctgactataaagaagagggactttcaaatgca
aactttgatctgaataataaattgggtattgttcattctcaacaagaagatgttaaaaaa
50  ctgattaagcgatatacagatttagtcatcgatcaattagaagatgatgactttgaaaaa
ggtaaaaaagaaaaagttaaaattgacggtgaaacaaaaaatttaaaacctatcactttа
aatataagtcgtgataaagcaaaaaaaatcactgtcgcagcattaaaaaaaagctaaaaat
gataaagaattac


55  Sequence 2748
MSQKIKVIVPIVLGLIILLGIAWGVYAFVTNTPKNAYLLSEKKTAINVKSYVDHRFSNEK
KFQKKLKDNSYVNTYNLHANASKEYLKDLGLPKTILDSSKITGTIGHDPKSNKGIMSVSP
KILDKDIGKFQWTANDSTQFFESPLFKKKYSVKNSELLETAAQIFDEDPSDYKEEGLSNA
NFDLNNKLGIVHSQQEDVKKLIKRYTDLVIDQLEDDDFEKGKKEKVKIDGETKNLKPITL

NISRDKAKKITVAALKKAKNDKELX

Sequence 2749
Contig_0813_pos_5275_4484,
putative peptide of unknown function
atgacaattcttgttcattcaaagcatcaaccgagcgaatacgcagcgattgcacatcag
ttgatggcgacaacacatgtgtgttgtgaacaagtagggttcattgaatcagtaaactat
gaaaatggggataactatcacttggtaatgagtggaaatgaattttgtggtaatgcgact
atgtcttacattcactatttaaaagaacgtttattgatacagcatcaacaatttcaatta
agagtttcgggatgttctcatcctgtagagtgtaaagttcattcgcaacattatgaagtg
actatgccaaaagtacatcaagttaaggaaagatttgtgaaattaggggaccaacagttt
aaagcatttgaaattagatacgatacatacattcattacgtgttgatgtgtgatggtgta
gatttagcaatgaaacagcgcgtggaagattttgtcagtgcgcaaacatggcatcaacaa
tttaaaacgattggcgtcatgcttttttcaacaagataaacaattcatatatccactgata
catatacctaaaatagatagcttaatctgggaaaatagctgtggttcaggagcggcttct
atcggtgtgttagttaattatctaacagatcatgatattcaagattacctagttaaccaa
cccggaggcagtattattgtctcatccagaaagtctggacaaaatgaataccaaacaacg
attaagtgtcaagtttcaactgtcgcaacaggacaagcatatatagaacaggagacaatg
acgcaaatatga

Sequence 2750
MTILVHSKHQPSEYAAIAHQLMATTHVCCEQVGFIESVNYENGDNYHLVMSGNEFCGNAT
MSYIHYLKERLLIQHQQFQLRVSGCSHPVECKVHSQHYEVTMPKVHQVKERFVKLGDQQF
KAFEIRYDTYIHYVLMCDGVDLAMKQRVEDFVSAQTWHQQFKTIGVMLFQQDKQFIYPLI
HIPKIDSLIWENSCGSGAASIGVLVNYLTDHDIQDYLVNQPGGSIIVSSRKSGQNEYQTT
IKCQVSTVATGQAYIEQETMTQI*

Sequence 2751
Contig_0813_pos_3664_2447,
putative peptide of unknown function
atggttggtagtggaccggtcgctattcaacttgctcgactatgtcatttacatggagaa
catatagttgatatggtgagtcgcgttcatgcatcaaccaaatctaagagagtctttgat
gcttatcaacgtgacggctttttttcagtaatgactcaaaatgatgcacatcagtgtttt
tcaggtaagtttacggttagacatttttttaaagatgttaaagatattactgaatattat
gacgtggtgattttagcatgtactgccgatgcgtatcgaccgatattacagcaattatct
aagtccacattaaagcgtattaagcaaatcatcttggtctcaccaacattaggatcacat
atccttgttaagcaattactatcagatgttcaatgtgaaggtgaagtgatttcattttcc
acttatcaggcgatacccgaatatttgataaagcacaaccacattgtgtcctaacaca
cgagttaaatcaaaattattcgtaggttcgactcaatctcagtctatgacgttgtgtaag
cttaagtctttatttgactatttgaatatagaattaacaacgatggacacaccactacat
gcggagatacataatagttcactttatgtacaccaccattgtttatgaatcaattttca
ttaaaggcggtatttgaagggacgaaagtaccagtatatgtatataagctatttccagag
ggtccaatcacaatgaccttaatacacgaaatgcgattaatgtggcaagaaatgatgatg
atattaaaaaaattaaaggtaccttcggtcaatcttctaaagtttatggtgaaagaaac
taccctatacgttatgagaccatgcgcgaagtagatattgaaaactttaaaaatttacca
gctattcatcaagagtatctactttatgtgcgatatacagcaatttttaatcgatccgttt
tctaatccggacgatcaaggtgcatattttgattttctgccgtaccatacaaacatgtt
gatactgatgaacaaggagtcatacatataccacgcatgccgagtgaagattattatcgt
actttgataattcaagcgattggaagagcattaaacgttgcaacaccgatgattgacaca
ttgttattacgttatgaaaatactgttaaacaatactgtgacacacatttacatcaacaa
ctatattccctaaattga

Sequence 2752
MVGSGPVAIQLARLCHLHGEHIVDMVSRVHASTKSKRVFDAYQRDGFFSVMTQNDAHQCF
SGKFTVRHFFKDVKDITEYYDVVILACTADAYRPILQQLSKSTLKRIKQIILVSPTLGSH
MLVKQLLSDVQCEGEVISFSTYLGDTRIFDKAQPHCVLTTRVKSKLFVGSTQSQSMTLCK
LKSLFDYINIELTTMDTPLHAEIHNSSLYVHPPLFMNQFSLKAVFEGTKVPVYVYILFPE
GPITNTLIHEMRLMWQEMMMILKKLKVPSVNLLKFMVKENYPIRYETMREVDIENIKNLP
AIHQEYLLYVRYTAILIDPFSNPDDQGAYFDFSAVPYKHVDTDEQGVIHIPRMPSEDYYR

672

TLIIQAIGRALNVATPMIDTLLLRYENTVKQYCDTHLHQQLYSLN*

Sequence 2753
Contig_0814_pos_338_709,
is similar to (with p-value 8.0e-51)
>gp:gp|D76414|D76414_4 Staphylococcus aureus gene for histid
yl-tRNA synthetase, ppGpp hydrolase, lytic enzyme, complete
cds. NID: g2580431.
atcactcgttccaatgataaatatgtgtccctagatgatagaaatataaagggtgaagca
tttattagtatacataatgatgcgcttgattcttcaaatgccaatggagtgactgttat
tggtttaaagacaagcaagaatcacttgcacaaactttaaattctgcaattcaaaagaag
gcattattgacaaatcgaggttctagacaacaaaattatcaagtgttgagacagacagat
ataccagcagtactgttagagttaggctatataagtaatcctactgatgaatcaatgatt
aatgatcaattacatagacaagtggttgaacaagctattgttgatggtttaaaacaatat
ttctcgtcctag

Sequence 2754
MTRSNDKYVSLDDRNIKGDAFISIHNDALDSSNANGVTVYWFKDKQESLAQTLNSAIQKK
ALLTNRGSRQQNYQVLRQTDIPAVLLELGYISNPTDESMINDQLHRQVVEQAIVDGLKQY
FSS*

Sequence 2755
Contig_0814_pos_1081_2355,
is similar to (with p-value 0.0e+00)
>sp:sp|O32422|SYH_STAAU HISTIDYL-TRNA SYNTHETASE (EC 6.1.1.2
1) (HISTIDINE--TRNA LIGASE) (HISRS). >gp:gp|D76414|D76414_5
Staphylococcus aureus gene for histidyl-tRNA synthetase, ppG
pp hydrolase, lytic enzyme, complete cds. NID: g2580431.
atgattaagatgccaagaggtacccaagatatcttgccgcaagattctgctaaatcgcgt
tacattgaaaatcgattacacacattaatggaattgtataattataaagaaataagaacg
ccaatttttgaaagtactgaactttttgcaagaggcgtgggggattctactgacgttgtt
caaaaggaaatgtatacatttaaagataaaggggatcgtagtttaacattacgtcctgaa
ggaactgcagccgttgtacgttcatatattgaacacaaaatgcaaggtgaacccaatcaa
cctatcaaactttactacaatggtcctatgtttagatacgaacgtaaacaaaaaggaaga
tatcgccaatttaaccaatttggtgtcgaagcaataggagcggaaaatcctagtattgat
gctgaaatactcgctatggttatgcatatatatgagtctttcggattaaagcatttaaag
ttagttatcaatagtattggtgatagtgaatcacgtaaagaatataacgaagcattagta
aaacatttgaacctgtgattgatacattttgttcagattgtcaatcaagattacacact
aatcctatgagaattttagattgtaaatcgatagagataaagaagcagtaaaaaatgca
ccccgtatcacagattatctcaataatgattctaaatcttattatgaacaagttaaatta
catcttgataatttgaacatatcttatgttgaagatcctaacttagttcgtgggttagat
tattatactcatactgcctttgaattaatgattgataatccagagtatgatggagctatc
actacattatgtggtggtggtcgatataatgggttgttacaattattagatggtccagat
gaaacaggtattgggtttgcactaagtattgaaagattattgatggcacttgatgaagaa
ggtatttcattagatgtaagtgaagattttgatttatttgttgtgacaatgggagaagat
gccgatcgttatgctgttaagttaatcaatgatttaagaagaaatggaataaaagtagat
aaggattatctaaacagaaaaattaaaggacaaatgaaacaagctgaccgtcttaatgct
aaatatacagtagtaattggagatcaagagcttgaaaataatgaaattggtgtgaaaaac
atgaatcaggcgaatcagaaaatgtacaattagacgaattggttaattattttacaagt
agaaaggaagtctaa

Sequence 2756
MIKMPRGTQDILPQDSAKWRYIENRLHTLMELYNYKEIRTPIFESTELFARGVGDSTDVV
QKEMYTFKDKGDRSLTLRPEGTAAVVRSYIEHKMQGEPNQPIKLYYNGPMFRYERKQKGR
YRQFNQFGVEAIGAENPSIDAEILAMVMHIYESFGLKHLKLVINSIGDSESRKEYNEALV
KHFEPVIDTFCSDCQSRLHTNPMRILDCKIDRDKEAVKNAPRITDYLNNDSKSYYEQVKL
HLDNLNISYVEDPNLVRGLDYYTHTAFELMIDNPEYDGAITTLCGGGRYNGLLQLLDGPD
ETGIGFALSIERLLMALDEEGISLDVSEDFDLFVVTMGEDADRYAVKLINDLRRNGIKVD
KDYLNRKIKGQMKQADRLNAKYTVVIGDQELENNEIGVKNMISGESENVQLDELVNYFKS

RKEV.*

Sequence 2757
Contig_0815_pos_2348_3526,
5    is similar to (with p-value 1.0e-92)
>sp:sp|P38494|RS1H_BACSU 30S RIBOSOMAL PROTEIN S1 HOMOLOG. >
gp:gp|U11687|BSU11687_5 Bacillus subtilis 168 jofA, jofB, Ms
sA homolog (jofC) and ribosomal protein S1 homolog (jofD) ge
nes, complete cds, and joeB gene, partial cds. NID: g533101.
10   >gp:gp|Z99115|BSUB0012_228 Bacillus subtilis complete genom
e (section 12 of 21): from 2195541 to 2409220. NID: g2634478
. >gp:gp|L47648|BACSERA_20 Bacillus subtilis phosphoglycerat
e dehydrogenase (serA), ypaA, ferredoxin (fer), ypbB, recS,
ypbD, ypbE, ypbF, ypbG, ypbH, glutamate dehydrogenase (ypcA)
15   , ypdA, ypdB, ypdC, spore cortex lytic enzyme (sleB), ypeB,
ypfA, ypfB, cytidine monophosphate kinase (cmk), ypfD, ypgA,
yphA, yphB, yphC, NAD+ dependent glycerol-3-phosphate dehyd
rogenase (glyc), yphE and yphF genes, complete cds. NID: g11
46195. >gp:gp|L47648|BACSERA_20 Bacillus subtilis phosphogly
20   cerate dehydrogenase (serA), ypaA, ferredoxin (fer), ypbB, r
ecS, ypbD, ypbE, ypbF, ypbG, ypbH, glutamate dehydrogenase (
ypcA), ypdA, ypdB, ypdC, spore cortex lytic enzyme (sleB), y
peB, ypfA, ypfB, cytidine monophosphate kinase (cmk), ypfD,
ypgA, yphA, yphB, yphC, NAD+ dependent glycerol-3-phosphate
25   dehydrogenase (glyc), yphE and yphF genes, complete cds. NID
: g1146195.
atgactgaagaattcaatgaatcaatgattaatgatattaaagaaggtgacaaagtcact
gttgaagttcaacaagtagaggataaacaagttgttgtgcatattaatggtggcaaattt
aatggaattattcctattagccagctttcaacacatcatatcgaaaaccctagtgaagtt
30   gtaaaagtcggtgatgaagtcgaagcatatgtcactaaaatcgagttcgacgaagaaaat
gatactggggcatacattttatcaaaaagacaacttgaaactgaaaaatcttatgaatat
ttacaagaaaaactagataacgatgaagtgattgaagctgaagttactgaagtagttaaa
ggtggtttagtcgttgacgttggtcaaagagggtttgtacctgcttctctaatttcaact
gatttcattgaagattttctgtattcgatggtcaaacaatccgtattaaagtggaagaa
35   cttgatcctgaaaacaatagagtcattttaagccgtaaagctgtggaacagttagaaaac
gacgctaaaaaagcttcaatattagattctttaaatgaaggcgatgttattgatggtaaa
gttgctcgattaactaactttggtgctttcattgatattggtggcgtagatggtttagtt
cacgtttctgaattatctcatgaacatgttcaaacaccagaagaagttgtgtcagtaggt
gaagcagtcaaagttaaagttaaatctgtagaaaaagattctgaacgtatttctttatct
40   attaaagacactttaccaacaccatttgaaaacattaaagggaaatttcacgaagatgat
gttattgaaggtactgtagtacgtttggcgaactttggcgcattcgtagaaattgctcca
tccgtccaaggtttagtgcatatttctgaaatcgatcataaacatatcggttctcctaac
gaagtattagaacctggacaacaagttaatgtaaaaatattaggtatcgatgaagataat
gaaagaatttcattatcaatcaaagcaacgttacctaaagaaaatgtcattgaaagtgac
45   gcatccacaactcaatcatatcttgaagatgataatgatgaagataaaccaacattaggc
gatgtttttggtgataaatttaaagaccttaagttttaa

Sequence 2758
MTEEFNESMINDIKEGDKVTVEVQQVEDKQVVVHINGGKFNGIIPISQLSTHHIENPSEV
50   VKVGDEVEAYVTKIEFDEENDTGAYILSKRQLETEKSYEYLQEKLDNDEVIEAEVTEVVK
GGLVVDVGQRGFVPASLISTDFIEDFSVFDGQTIRIKVEELDPENNRVILSRKAVEQLEN
DAKKASILDSLNEGDVIDGKVARLTNFGAFIDIGGVDGLVHVSELSHEHVQTPEEVVSVG
EAVKVKVKSVEKDSERISLSIKDTLPTPFENIKGKFHEDDVIEGTVVRLANFGAFVEIAP
SVQGLVHISEIDHKHIGSPNEVLEPGQQVNVKILGIDEDNERISLSIKATLPKENVIESD
55   ASTTQSYLEDDNDEDKPTLGDVFGDKFKDLKF*

Sequence 2759
Contig_0815_pos_4093_5052,
is similar to (with p-value 0.0e+00)

>sp:sp|P50743|YPHC_BACSU HYPOTHETICAL 48.8 KD GTP-BINDING PR
OTEIN IN CMK-GPSA INTERGENIC REGION. >gp:gp|Z99115|BSUB0012_
224 Bacillus subtilis complete genome (section 12 of 21): fr
om 2195541 to 2409220. NID: g2634478. >gp:gp|L47648|BACSERA_
24 Bacillus subtilis phosphoglycerate dehydrogenase (serA),
ypaA, ferredoxin (fer), ypbB, recS, ypbD, ypbE, ypbF, ypbG,
ypbH, glutamate dehydrogenase (ypcA), ypdA, ypdB, ypdC, spor
e cortex lytic enzyme (sleB), ypeB, ypfA, ypfB, cytidine mon
ophosphate kinase (cmk), ypfD, ypgA, yphA, yphB, yphC, NAD+
dependent glycerol-3-phosphate dehydrogenase (glyc), yphE an
d yphF genes, complete cds. NID: g1146195. >gp:gp|L47648|BAC
SERA_24 Bacillus subtilis phosphoglycerate dehydrogenase (se
rA), ypaA, ferredoxin (fer), ypbB, recS, ypbD, ypbE, ypbF, y
pbG, ypbH, glutamate dehydrogenase (ypcA), ypdA, ypdB, ypdC,
spore cortex lytic enzyme (sleB), ypeB, ypfA, ypfB, cytidin
e monophosphate kinase (cmk), ypfD, ypgA, yphA, yphB, yphC,
NAD+ dependent glycerol-3-phosphate dehydrogenase (glyc), yp
hE and yphF genes, complete cds. NID: g1146195.
gtgaataaagttgataatcttgaaatgcgtaatgatatctatgatttctattctttaggc
tttggagatccatatcctatttctggttcacatggattaggacttggagatttgctagat
gcagttgttgaaaactttaataaagaatcagaagatccttatgacgaagatacgatacgt
ctttctatcatcggtagacctaatgttggtaaatctagcttggtcaatgctattttaggc
gaagaacgtgttattgtgtctaatgttgctggtacaactcgagatgccattgataccgag
tactcttatgatggacaagattatgtattgattgatactgctggaatgagaaaaaaaggt
aaggtgtatgaatcgactgaaaaatattctgtattacgtgcattaaaagcgattgagcgt
tcagaagtagtattagtagttatcgatgctgaacaaggtataattgaacaagataaacgt
gtagctggctatgcacatgaggaaggtaaagctattgtcattgtagtaaataaatgggat
acagttgaaaaagatagtaagacaatgaaaaaattcactgatgatgttagaaatgaattt
caatttttagattatgctcaaatcgcgttcgtatcagcaaaagaagggctaagattaaaa
acattattcccttatatcaatcaagccagtgaaaatcataaaaagcgtgtccaaagttct
acactaaatgaagttgttactgatgccatctctatgaatccaacacctactgacaaaggt
agaagact.taatgtattctatacaactcaggttgcaattgaaccaccgacatttgtagta
tttgtcaatgatgttgaattaatgcatttttcttataggagatatttagaaaatcaaata
cgtaatgcttttggttttgaaggaacacctattcatattattccaagaaaaagaaattaa

Sequence 2760
VNKVDNLEMRNDIYDFYSLGFGDPYPISGSHGLGLGDLLDAVVENFNKESEDPYDEDTIR
LSIIGRPNVGKSSLVNAILGEERVIVSNVAGTTRDAIDTEYSYDGQDYVLIDTAGMRKKG
KVYESTEKYSVLRALKAIERSEVVLVVIDAEQGIIEQDKRVAGYAHEEGKAIVIVVNKWD
TVEKDSKTMKKFTDDVRNEFQFLDYAQIAFVSAKEGLRLKTLFPYINQASENHKKRVQSS
TLNEVVTDAISMNPTPTDKGRRLNVFYTTQVAIEPPTFVVFVNDVELMHFSYRRYLENQI
RNAFGFEGTPIHIIPRKRN*

Sequence 2761
Contig_0815_pos_5073_6071,
is similar to (with p-value 0.0e+00)
>gp:gp|Z99115|BSUB0012_223 Bacillus subtilis complete genome
(section 12 of 21): from 2195541 to 2409220. NID: g2634478.
>gp:gp|L47648|BACSERA_25 Bacillus subtilis phosphoglycerate
dehydrogenase (serA), ypaA, ferredoxin (fer), ypbB, recS, y
pbD, ypbE, ypbF, ypbG, ypbH, glutamate dehydrogenase (ypcA),
ypdA, ypdB, ypdC, spore cortex lytic enzyme (sleB), ypeB, y
pfA, ypfB, cytidine monophosphate kinase (cmk), ypfD, ypgA,
yphA, yphB, yphC, NAD+ dependent glycerol-3-phosphate dehydr
ogenase (glyc), yphE and yphF genes, complete cds. NID: g114
6195. >gp:gp|L47648|BACSERA_25 Bacillus subtilis phosphoglyc
erate dehydrogenase (serA), ypaA, ferredoxin (fer), ypbB, re
cS, ypbD, ypbE, ypbF, ypbG, ypbH, glutamate dehydrogenase (y

675

pcA), ypdA, ypdB, ypdC, spore cortex lytic enzyme (sleB), yp
eB, ypfA, ypfB, cytidine monophosphate kinase (cmk), ypfD, y
pgA, yphA, yphB, yphC, NAD+ dependent glycerol-3-phosphate d
ehydrogenase (glyc), yphE and yphF genes, complete cds. NID:
g1146195.
atgagaaaaattacagttttttggtatgggtagttttggtactgcattagctaatgtatta
gctcaaaatggtcatgatgtttttaatgtggggcaaaaatgtagagaatgtagatgaactt
aacacacatcatatgaacaaaaattatcttaaagatgctaaattagattcatctataaaa
gcaactgtagatttaaataaggcagtgcaattttcagatatctatcttatggcactacct
acaaaagcaattagagaagtatcaaaagatatcgatcaattactcacatctaaaaagact
tttattcatgttgctaaaggcattgaaaacgatacatttaagcgcgtatctgaaatgatt
gaggactctatctcttcagaacataatggaggaatcggcgtcttatcaggtccaagtcat
gctgaagaagttgttataaaacaacctacaactgtagctgcatcatctaaagataataat
gtgagcaaacttattcaagatttatttatgaacgactatttacgtgtttacacaaataat
gatttagtaggtgtagaattaggtggtgctttaaaaaatattatagctatagctagtggt
atcgttgccggcatgggttacggtgataatgcaaaagcagctttaatgacacgaggttta
gccgaaatcagtcgacttggtgagaaacttggtgcagatccaatgactttcttaggtcta
ggtggcataggtgacttgatcgtaacttgtacgtccacacattcacgaaattacacactt
gggtttaaattagggcaaggcaaaacagcagaagaagctttaaaagagatgaaaatggtg
gttgaaggtatttatacaactaaatcagtatatcatcttgctcaacaagaaggagtagag
atgcctatcactaacgcattatatgaagttttatttgaagatgtccctgtaagtaaaagt
gttagaacacttatggaaagagacaaaaaagcagaataa

Sequence 2762
MRKITVFGMGSFGTALANVLAQNGHDVLMWGKNVENVDELNTHHMNKNYLKDAKLDSSIK
ATVDLNKAVQFSDIYLMALPTKAIREVSKDIDQLLTSKKTFIHVAKGIENDTFKRVSEMI
EDSISSEHNGGIGVLSGPSHAEEVVIKQPTTVAASSKDNNVSKLIQDLFMNDYLRVYTNN
DLVGVELGGALKNIIAIASGIVAGMGYGDNAKAALMTRGLAEISRLGEKLGADPMTFLGL
GGIGDLIVTCTSTHSRNYTLGFKLGQGKTAEEALKEMKMVVEGIYTTKSVYHLAQQEGVE
MPITNALYEVLFEDVPVSKSVRTLMERDKKAE*

Sequence 2763
Contig_0815_pos_0_368,
is similar to (with p-value 2.0e-21)
>sp:sp|P42086|PBUX_BACSU XANTHINE PERMEASE. >pir:pir|S51310|
S51310 xanthine permease - Bacillus subtilis >gp:gp|L77246|B
ACYACA_3 Bacillus subtilis (YAC10-9 clone) DNA region betwee
n the serA and kdg loci. NID: g1256615. >gp:gp|Z99115|BSUB00
12_147 Bacillus subtilis complete genome (section 12 of 21):
 from 2195541 to 2409220. NID: g2634478. >gp:gp|X83878|BSXPT
PBUX_2 B.subtilis xpt and pbuX genes. NID: g633168.
atgtgcggggtagcgacatttcttcaagcaaataaagtcacagggactggattaccgatt
gtactaggatgtacgtttactgccgttgcacctatgatactcatcggtcaaacgaaagga
cttgatgtttatatggttcgctttaatatccggtatcttagttgttttaattgcacct
ttttctcttatttagttaaattctttccacctgttgtaacaggaagtgttgtgacaatt
attggaatcaatttaatgccagttgcaatgaattacttggcaggtggtgaaggagcgaaa
aactatggcgatactaagaatttaatattaggtggtgttacactactcattattcttatt
tGATTTAT

Sequence 2764
MCGVATFLQANKVTGTGLPIVLGCTFTAVAPMILIGQTKGLDVLYGSLLISGILVVLIAP
FFSYLVKFFPPVVTGSVVTIIGINLMPVAMNYLAGGEGAKNYGDTKNLILGGVTLLIILI
*FX

Sequence 2765
Contig_0818_pos_4072_3692,
putative peptide of unknown function
gtgtttattttgttaacatttggattttatgtattttttgctggccataataatccaggt
ggtggctttattggtggcttgattttagctcggcatttatcttaatgtttcttgcctttt

```
gatgtaaatgaagtgttgaaaagcttgcctattgattttaaaaaattaatgattataggt
tcactcatatctgttgcaactgcatcagtgcctatgttttttgggaagccattttttatat
caaactgaagcaaatgtaacatttccattactaggacatgttcatgttactactgtgact
ttatttgagcttggcatcttattaacagtagtaggtgtgattgttacagttatgctatct
ataagtgggggtagatcatga
```

Sequence 2766
VFILLTFGFYVFFAGHNNPGGGFIGGLIFSSAFILMFLAFDVNEVLKSLPIDFKKLMIIG
SLISVATASVPMFFGKPFLYQTEANVTFPLLGHVHVTTVTLFELGILLTVVGVIVTVMLS
ISGGRS*

Sequence 2767
Contig_0818_pos_3674_3351,
is similar to (with p-value 3.0e-20)
>gp:gp|AB015981|AB015981_4 Staphylococcus aureus genes for O
rfA, MnhA, MnhB, MnhC, MnhD, MnhE, MnhF and MnhG, complete c
ds. NID: g4001723.

```
gtgataggattttttagtgtttattggaacttatatgattttatctattaatttaattcgt
attgttattggtatttctatttatacacacgccggtaatttaattattatgagtatgggg
aaatatggacctcatatgtctgaaccgctaattcaaggtcatgctcaaaactttgttgat
cctttattacaagctatcgttttaacagctattgtgattggatttggtatgactgcgttt
ttattggtgttaatatatagaacttacagagtaactaaagaggatgaaataagtgcattg
aaaggtgatgaagatgatgagtaa
```

Sequence 2768
VIGFLVFTGTYMILSINLIRIVIGISIYTHAGNLIIMSMGKYGPHMSEPLIQGHAQNFVD
PLLQAIVLTAIVIGFGMTAFLLVLIYRTYRVTKEDEISALKGDEDDE*

Sequence 2769
Contig_0818_pos_3334_1859,
is similar to (with p-value 2.0e-62)
>gp:gp|AB015981|AB015981_5 Staphylococcus aureus genes for O
rfA, MnhA, MnhB, MnhC, MnhD, MnhE, MnhF and MnhG, complete c
ds. NID: g4001723.

```
atgttgttgccttttgtatgtgctttaattttagtcttcactaaaaataaaaatcgtatt
tcgaaaatcctatccattacaactatgattgttaatacaatgatttcaattgctttactt
atttatgtcgttaatcataaaccgataacacttgattttgggggatggaaagcacctttc
ggcattcaatttctaggtgattcactgagtctgcttatggtgtcagtatcatcttttgtt
gttacgctaataatggcatacggctttggtagaggggagaagcgagtcaatcgatttcac
cttcctacatttattcttttattaacagtaggtgttattggttcgttttttaacttctgat
ttatttaacctatacgtgatgtttgaaattatgcttcttgcttcgtttgtacttgttaca
ttaggacaatctgttgaacaattacgtgcagcgatagtatatgttgttctgaatatttta
ggttcgtggttgcttttattaggaattggcatgttatataagacagtcggaacacttaat
ttctcacatttagcgatgcgattgaatcatatggaaaataaccaaacaataacgatgata
tctttagtatttctagttgctttttagttcaaaggcagcactagtgattttcatgtggtta
cctaaagcatatgcagtgcttaatacggaacttgccgcgttatttgcagcattgatgaca
aaaagttggagcttatgcgcttattcgttttttttactttactattcgaccatcatccaagc
gtcacgcatacattgctcgtgtttatggcttgtatcacaatgattatcggtgcatttggt
gtcatcgcttacaaagatattaagaaaattgcggcttatcaagttattttgtctattgga
ttcattattttaggtttaggttctcatactatatcaggtgtaaatggtgctatcttctat
ttagcgaatgatattatcgttaagacattattgtttttgtaattggtagtcttgtttat
atgtcaggctatcgaaattatcagtatttaagtggactggcaaaaagagaaccattcttt
ggtgttgcatttgtcgtggtaattttttgctataggtggcgtacctccttttagtggcttt
ccgggtaaagtcttaatattccaagggggctattacaaatggtaattatattggtttagca
cttatgattgtgacaagtttaattgctatgtatagtcttttttagagtgatgtttataatg
tattttggtgatgctgacggagaacaagtacaatttagaccactacctatttatcgtaaa
ggtttacttagtgtttttagttgtagtggtattagcgatgggtattgcagcccctgttgtt
ctgaaagtaacagaggatgcaacaaatcttaatatgaaagaagatgtctttcaaaagaat
gtaaatacacatttgaaggaggttaatcataagtga
```

Sequence 2770
MLLPFVCALILVFTKNKNRISKILSITTMIVNTMISIALLIYVVNHKPITLDFGGWKAPF
GIQFLGDSLSLLMVSVSSFVVTLIMAYGFGRGEKRVNRFHLPTFILLLTVGVIGSFLTSD
5 LFNLYVMFEIMLLASFVLVTLGQSVEQLRAAIVYVVLNILGSWLLLLGIGMLYKTVGTLN
FSHLAMRLNHMENNQTITMISLVFLVAFSSKAALVIFMWLPKAYAVLNTELAALFAALMT
KVGAYALIRFFTLLFDHHPSVTHTLLVFMACITMIIGAFGVIAYKDIKKIAAYQVILSIG
FIILGLGSHTISGVNGAIFYLANDIIVKTLLFFVIGSLVYMSGYRNYQYLSGLAKREPFF
GVAFVVVIFAIGGVPPFSGFPGKVLIFQGAITNGNYIGLALMIVTSLIAMYSLFRVMFIM
10 YFGDADGEQVQFRPLPIYRKGLLSVLVVVVLAMGIAAPVVLKVTEDATNLNMKEDVFQKN
VNTHLKEVNHK*

Sequence 2771
Contig_0818_pos_1751_1380,
15 putative peptide of unknown function
gtgatttatattctgcatcgctttttttggtgaagaattttatttgaaaaagatatgggtg
gctattaaattttttagctgtatacctataccagcttattacttctagtataagtaccata
aattacatcttatttaagacgaatgaagttaatccaggtttactcacatatgaaacttca
ttaaaaagtaattgggctattacttttttttaacgattttaattattattactccaggatcg
20 acagttattcgaatttctaaaaatactaataaattttttattcacagtattgatgtgtca
gaaaaagataaagaaatcttctaaaaagtattaagcagtatgaggatttaattttggag
gtgacacgatga

Sequence 2772
25 VIYILHRFFGEEFYLKKIWVAIKFLAVYLYQLITSSISTINYILFKTNEVNPGLLTYETS
LKSNWAITFLTILIIITPGSTVIRISKNTNKFFIHSIDVSEKDKENLLKSIKQYEDLILE
VTR*

Sequence 2773
30 Contig_0819_pos_5530_6732,
is similar to (with p-value 0.0e+00)
>sp:sp|Q53634|MENE_STAAU O-SUCCINYLBENZOIC ACID--COA LIGASE
(EC 6.2.1.26) (OSB-COA SYNTHETASE) (O-SUCCINYLBENZOATE-COA S
YNTHASE). >gp:gp|U51132|SAU51132_1 Staphylococcus aureus o-s
35 uccinylbenzoic acid CoA ligase (mene), and o-succinylbenzoic
acid synthetase (menc) genes, complete cds. NID: gl255258.
atgataaatacacgtttaacgcgacatgagatgataaatcaaatgaattcagtcgacata
gcaacgattgtacacacgttgcctttagaattagaagggtttaatttatatcatttaat
gatttaacacaattagataaacatgatgtttcaggttacaaatttaatttagaatcgatt
40 gcatcaattatgtttacgtctggaacgacgggacctcaaaaagctgtgcctcaaacgttt
aataatcatttagccagtgctaaaggctgtaaacaaagtttaggattcgaacaaatact
gtgtggctttcggtcttacctatatatcatattctgggctcagtgttattttgcgcgca
gtgatagaaggattcactgtcagacttgttaaaaagtttcaaactgatgatatgttaaca
caaataaagacttatccaatcacccatatgtcccttgttccacaaacgttaaagtggtta
45 atggatgcaggattgactcaaccattttcttttgaaaaaaattctgctaggtggtgctaaa
ttatcaccacaattaattgagcaagcattgacttatcgtttacctgtatataattcttttt
ggtatgacagaaacttgctctcagtttctaacagcctcacctcaaatgctcaaagaacgt
ttcgatactgttggaaaaccaagtgaaaatgtcgaagtgaaaataaaaaaatcccaacgca
tatggacatggagagttattaattaaaggtgaaaatgtgatgaatggttatttatatccc
50 aaatatttaaaagacacatttgataatgatgggtattttcaaactggagatatagctgaa
atagatgatgaaggttacgtcataatatatgatcggcgcaaagatttgattataagtggt
ggagagaatatttatccttaccaaattgaaacaatcgcaaaagactttgaaggcattgaa
gatgccgtatgtgtaggaatatcagatgatacttggggtcaagtaccaatattatattat
gtgacaaatcaagatattaatcaaactgaattaatagaacattttgagaatcatttagct
55 agatataaaattcctaaaaaatattatcaggtcaaatctttaccttatacatcgacaggt
aaattacaacgtaaaaaggtcaaaagtgaagacttgaatgagggaaagaataatgaaagt
taa

Sequence 2774

MINTRLTRHEMINQMNSVDIATIVHTLPLELEGFNLYHFNDLTQLDKHDVSGYKFNLESI
ASIMFTSGTTGPQKAVPQTFNNHLASAKGCKQSLGFEQNTVWLSVLPIYHISGLSVILRA
VIEGFTVRLVKKFQTDDMLTQIKTYPITHMSLVPQTLKWLMDAGLTQPFSLEKILLGGAK
LSPQLIEQALTYRLPVYNSFGMTETCSQFLTASPQMLKERFDTVGKPSENVEVKIKNPNA
5   YGHGELLIKGENVMNGYLYPKYLKDTFDNDGYFQTGDIAEIDDEGYVIIYDRRKDLIISG
GENIYPYQIETIAKDFEGIEDAVCVGISDDTWGQVPILYYVTNQDINQTELIEHFENHLA
RYKIPKKYYQVKSLPYTSTGKLQRKKVKSEDLNEGKNNES*

Sequence 2775
10   Contig_0819_pos_6791_0,
is similar to (with p-value 6.0e-43)
>gp:gp|U51132|SAU51132_2 Staphylococcus aureus o-succinylben
zoic acid CoA ligase (mene), and o-succinylbenzoic acid synt
hetase (menc) genes, complete cds. NID: g1255258.
15   gtgaagttaacgcatcgagaaagtttgtttacggaaatagtaacttatagtggagaaact
tattatggagaatgtaatgcatttttaactaattggtatgataaagaaacaatactcaca
gtcgtaaacagattaagacagtggataccgcaagtacttcataaagatatgacatcttt
gattcatggttaccttacctaaatcaaatgaatgatgcgccagctgctagatcaatggtt
gtcatggctgtttatcaaatgtataacgacttgcatgattttgaagtacaatacggtgct
20   acagttagtggcttaactaatagtcaaattgaaacattattagaaacaagaccgaaacgt
ataaaacttaaatggtcaacatcactcatcaaagatcttgaaactatacgtttattaaat
tttgattgtgatattgctatagatgcaaatgaatcattaacaaagccatcattttttacaa
ttagccaacgtaaatacatcagatattatatatattgaagaaccttttaaaattct

25   Sequence 2776
VKLTHRESLFTEIVTYSGETYYGECNAFLTNWYDKETILTVVNRLRQWIPQVLHKDMTSF
DSWLPYLNQMNDAPAARSMVVMAVYQMYNDLHDFEVQYGATVSGLTNSQIETLLETRPKR
IKLKWSTSLIKDLETIRLLNFDCDIAIDANESLTKPSFLQLANVNTSDIIYIEEPFKIL

30   Sequence 2777
Contig_0819_pos_4059_3394,
putative peptide of unknown function
atgcaacaagaaacgacatcatggtacaaacaagaatggtttatagtttttatcactttta
ttcattttttccactaggtttatttctcatgtggaaatttagcaagtggccatctattgca
35   agaacaatcattactgttgcaatttcagttatcgtattagcaagcattacctattatggt
aatctacaaatgattgtaccagcaacatcaaattcaaataacgaaactaaagaaactaca
gagaataatgtaaatgataaagacgagcgaaatcataaaactgcagtagaagaaacaaaa
actaattatgactccaccaaagaaaatactaaagaacctggaaaagaaaatgaatctgca
acacgattggagaactctgcgcttgaaaaggcaaagtcatattatgatgattttcacatg
40   tctaaactaggaatttatgatattttaacatctgaatatggagaaaaatttgataaagaa
gatgcacaatatgctatagatcatctagaggctgattatgaaaagaatgcacttgagaaa
gcaaaatcatatgccaaagatatgcatatgtctaatgactcaatttacgatcttttggtg
tctaactacggtgaaaaatttacagaatcagaagcaaaatatgctattgagcatttggat
aattaa
45

Sequence 2778
MQQETTSWYKQEWFIVLSLLFIFPLGLFLMWKFSKWPSIARTIITVAISVIVLASITYYG
NLQMIVPATSNSNNETKETTENNVNDKDERNHKTAVEETKTNYDSTKENTKEPGKENESA
TRLENSALEKAKSYYDDFHMSKLGIYDILTSEYGEKFDKEDAQYAIDHLEADYEKNALEK
50   AKSYAKDMHMSNDSIYDLLVSNYGEKFTESEAKYAIEHLDN*

Sequence 2779
Contig_0819_pos_2417_1683,
putative peptide of unknown function
55   atgagaaaattatttctatcaattttatctattataataatttcaagttttttgtgttgca
acagggtttcaaaacgttaatgctgcaaataatgaggtgagcaaacctcaaagtaatgtg
gatagtaaaactaaacaaaatattattaaaaaaataaaaaaatctaatgcttataaaaag
cacgctaattcaacttctattgactcaataaaagatgacgatattattgttcatttggat
aaaggtaaaaacaccaatgtatattctattaactatgtatttggtaaaaaaattagctaaa

atgaatgataaccttgcaatgattgaattaaaatataacgaagcaaacaatgaagttttc
tacaatcatatgatgtattcaaaatttgttaaatataataataaagaatatattaatatg
aaaggcatttaaatggcaaaccttattatgaatttaatatagatcagaaaggtcttac
tacgataaaaactttaaacatacctctaaagatgaaatcgaaaaagattctgctaaaaac
5   ttaccacctaaagagcgtggttggtgtgaatgggcagtaggagctttatgtggtaccggt
ggagctgcaggttgttgggcaactgctacagctttaggtattactactggttggggaggc
ttttcattagctacaatttgtggtctgataagctctctaggctgtactggtgcaaccaac
tatatttgtaaataa

10 Sequence 2780
MRKLFLSILSIIIISSFCVATGFQNVNAANNEVSKPQSNVDSKTKQNIIKKIKKSNAYKK
HANSTSIDSIKDDDIIVHLDKGKNTNVYSINYVFGKKLAKMNDNLAMIELKYNEANNEVF
YNHMMYSKFVKYNNKEYINMKGILNGKPYYEFNIDQKGHYYDKNFKHTSKDEIEKDSAKN
LPPKERGWCEWAVGALCGTGGAAGCWATATALGITTGWGGFSLATICGLISSLGCTGATN
15 YICK*

Sequence 2781
Contig_0822_pos_6711_5857,
is similar to (with p-value 2.0e-42)
20 >gp:gp|D78193|BACGNTZA_33 Bacillus subtilis 36kb sequence be
tween gntZ and trnY genes encoding 34 ORFs. NID: g1064780. >
gp:gp|Z99124|BSUB0021_145 Bacillus subtilis complete genome
(section 21 of 21): from 3999281 to 4214814. NID: g2636442.
atgaagtggcttaaacaactacaatcccttcacacgaaactcgttattgtttatgtacta
25 ctcattattattggtatgcaaatcatcggtttgtattttacgaatagtttagaaaaggaa
ttactcgataacttcaagaagaacataacacaatatgcgaagcaattagacgtcaatatt
gaaaaggtttataaagataaagataaaggttcagtcaacgctcaaaaggatatccaagac
cttttgaatgaatatgcgaatcgccaagaaataggagaaatacgctttattgataaagac
caaattatcatggcaacaaccaagcagtctaaccgtggtcttatcaatcaaaaggttaac
30 gacggttcagttcaaaaggcgctctccttagggcaaacgaatgatcatatggttcttaag
gattacggaagtggtaaagagcgtgtttgggtatataatataccggttaaagttgataaa
cagacaatcggtgatatatacatagaatcgaaaattaatgatgtatacaatcagctgaac
aacattaatcagatattcatcgtagggacagcgatatcactattcattacagtaatacta
ggattcttcattgcacgaacgattactaagccgataaccgatatgcgtaaccaaaccgtt
35 gagatgtctaaaggtaactacacgcaacgagtgaagatatacggtaacgatgaaatcggt
gagctcgcacttgccttcaataacttatcgaaacgtgtccaagaagcacaagcgaataca
gaaagtgagaaacgtcgcctagattctgttatcacacataattacttgttgaatgtatta
cattttgattgttaa

40 Sequence 2782
MKWLKQLQSLHTKLVIVYVLLIIIGMQIIGLYFTNSLEKELLDNFKKNITQYAKQLDVNI
EKVYKDKDKGSVNAQKDIQDLLNEYANRQEIGEIRFIDKDQIIMATTKQSNRGLINQKVN
DGSVQKALSLGQTNDHMVLKDYGSGKERVWVYNIPVKVDKQTIGDIYIESKINDVYNQLN
NINQIFIVGTAISLFITVILGFFIARTITKPITDMRNQTVEMSKGNYTQRVKIYGNDEIG
45 ELALAFNNLSKRVQEAQANTESEKRRLDSVITHNYLLNVLHFDC*

Sequence 2783
Contig_0824_pos_5618_5929,
putative peptide of unknown function
50 atgcacataaattcaaaattagctaaaatgccgcgtagaattgatggcaagttattctcg
gttagtcatacatcagtaagtaaagcgtttagaaaagcaaaagaagtgataggattaaac
gataataatataactccctattcactcagacatacgcacacatcttacttactatctaaa
ggcataccaatcgagtatataagtaaacgtttaggtcacgctactatatcacaaacgtta
gacacgtattcacatttattagaagaacataaaaaagagcaaggccaacgtgtcagagaa
55 atattctcttga

Sequence 2784
MHINSKLAKMPRRIDGKLFSVSHTSVSKAFRKAKEVIGLNDNNITPYSLRHTHTSYLLSK
GIPIEYISKRLGHATISQTLDTYSHLLEEHKKEQGQRVREIFS*

Sequence 2785
Contig_0824_pos_4770_3757,
putative peptide of unknown function
atggaacgattttgttgtgtaaatcaaattaactatattcaaatgaatccgttagaagcc
aaatttaaaacgagcgctctaagatcatggaaaactgatcaggcagatgctcataagctt
gcttgtttaggaccgacgcttaaacaaacagacaacttacctatacatgagttaatattc
tttgaattaagagaacgcgtccgtttttcatctagaaatcgagaatgaacaaaatcgactt
aaatttcagatccttgaattactccatcaaacattccctggtttagaaagattatttagt
agtcgatattcaatcattgcactcaacatcgcagaaatctttactcatccagacatggtt
cttgatatcgacaaggaggtactgattacacatatattcaattctacagataagggaatg
tcaatggataaagctacaaaatatgcacttcaattaagggtgattgctcaagaaagctat
cctaatgtcgatagacattcctttctagtcgaaaaattacgcttacttattcaacaatta
aaacaatctattcatcatctcaaacaattagatgatgccatgattcaattagcacaacaa
ctcgattattttgaaaatattcattcgatacctggtattggtaagctaagcacagctatg
attattggggagattggtgatattaagcgatttaaatcaaataaacaactcaatgctttt
gttggcattgatatcaaacgatatcaatcaggtcatacacactgtagagataccatcaac
aagcgtggtaataaaaaagcgagaaaactttattttgggtgattatgaatataataaga
gggcagcatcattatgacaatcatgtcgtcgattattactacaaactaagaaagcagcct
aatgagaaacctcataagactgccatcattgcttgtataaatcgattattaaaaacaatt
cattatcttgtaatgaatcataaattgtacgattatcaaatgtcaccacattag

Sequence 2786
MERFCCVNQINYIQMNPLEAKFKTSALRSWKTDQADAHKLACLGPTLKQTDNLPIHELIF
FELRERVRFHLEIENEQNRLKFQILELLHQTFPGLERLFSSRYSIIALNIAEIFTHPDMV
LDIDKEVLITHIFNSTDKGMSMDKATKYALQLRVIAQESYPNVDRHSFLVEKLRLLIQQL
KQSIHHLKQLDDAMIQLAQQLDYFENIHSIPGIGKLSTAMIIGEIGDIKRFKSNKQLNAF
VGIDIKRYQSGHTHCRDTINKRGNKKARKLLFWVIMNIIRGQHHYDNHVVDYYYKLRKQP
NEKPHKTAIIACINRLLKTIHYLVMNHKLYDYQMSPH*

Sequence 2787
Contig_0824_pos_2942_2442,
is similar to (with p-value 7.0e-38)
>gp:gp!AF082668|AF082668_1 Streptococcus pyogenes CsrR (csrR
) and CsrS (csrS) genes, complete cds. NID: g3599370.
atgcttccaaacataaatggtctagaaatttgtagacaaattcgtcaaaaaacaactact
ccaattatcatcattactgcaaaaagcgagacatatgataaagtagctgggttggactat
ggggcagatgactacattgtaaaacccttttgatatagaagaattgctcgcaagaataaga
gcggtattgcgcagacagccagataaagatgttttagatatcaatggtattatcattgat
aaagatgcctttaaagttactgttaatggccatcaattagaattaactaaaacagaatac
gatttattatatgtttttagctgaaaatcgtaaccacgtcatgcagcgtgaacaaattctc
gatcacgtatggggtataatagtgaagtagaaacgaatgtcgttgatgtttacattcgt
tatttacgtaataaactcaaaccttttaataaagaaaaatccatagaaacagtacgtggc
gtagggtatgtgattcgatga

Sequence 2788
MLPNINGLEICRQIRQKTTTPIIIITAKSETYDKVAGLDYGADDYIVKPFDIEELLARIR
AVLRRQPDKDVLDINGIIIDKDAFKVTVNGHQLELTKTEYDLLYVLAENRNHVMQREQIL
DHVWGYNSEVETNVVDVYIRYLRNKLKPFNKEKSIETVRGVGYVIR*

Sequence 2789
Contig_0824_pos_1929_1075,
is similar to (with p-value 2.0e-42)
>gp:gp|U81166|LLU81166_1 Lactococcus lactis subsp. cremoris
MG1363 histidine kinase (llkinA) gene, complete cds. NID: g2
182834.
gtgagttatatctttttcttcgcaaattactaaaccgatagttacaatgtccaataaaatg
aatcaaattagaagagatggttttcaaaataaacttgaattaactacaaattatgaagaa
acagataatttaattgatacttttaatgaaatgatgtatcaaatagaagaatcttttaat

```
cagcaacgtcaatttgtcgaggatgcttcacacgaattaagaacgccactgcagattatt
caaggtcatctaaatttaatccaacgttgggggaaaaaagatccagcagttttggaagaa
tctttgaatatttcaattgaagaagtgaatcgaataacaaaacttgtcgaagaactactt
```
5
```
ttacttaccaaagatagagtcaatcataatgttttggaatgtgaaaatgtagacgtaaat
agcgagattcaatcacgtgtgaagtcactgcaacacctacatccagattatacttttgaa
acacatcttgctactaagcctatccaattaaaaattaaccgtcatcagtttgaacaactc
ttactcatatttattgataatgcaatgaaatacgacactgaacataagcacattaaaatt
gttactcaactaaaaaataaaatgattatgattgatattactgatcatggtatgggtata
ccaaaagctgacttagaatttatctttgatagattttatcgtgtagataaatcacgtgct
```
10
```
cgtagtcaaggaggcaatggattaggactatcaatagcagaaaaaattgtgcaacttaac
ggtggtatgattcaagtagaaagtgaactacaaaagtacacgactttcaaaatcagtttt
ccagtactaaactaa
```

15
Sequence 2790

```
VSYIFSSQITKPIVTMSNKMNQIRRDGFQNKLELTTNYEETDNLIDTFNEMMYQIEESFN
QQRQFVEDASHELRTPLQIIQGHLNLIQRWGKKDPAVLEESLNISIEEVNRITKLVEELL
LLTKDRVNHNVLECENVDVNSEIQSRVKSLQHLHPDYTFETHLATKPIQLKINRHQFEQL
LLIFIDNAMKYDTEHKHIKIVTQLKNKMIMIDITDHGMGIPKADLEFIFDRFYRVDKSRA
RSQGGNGLGLSIAEKIVQLNGGMIQVESELQKYTTFKISFPVLN*
```

20

Sequence 2791
Contig_0824_pos_0_597,
is similar to (with p-value 8.0e-29)
>pir:pir|S25295|A32879 oxoglutarate dehydrogenase (lipoamide
25) (EC 1.2.4.2) - Bacillus subtilis
```
atgctagatttgtatgatgattatctacaaaatccatcatccgtacctgaagatttacaa
gtcttgttcagtacaattaaaacaggtgaagctcatatcgaagctaaacctaccactgat
gggggtggttcacaagcgggagatagcacaattaaacgtgttatgcgcttaatcgataat
attcgtcaatacggacatttaaaagcagatatttatccagtaaatcctccagagcgtcaa
```
30
```
aatgttcctaaattggaaatcgaagattttgatttagataaagaaactttggaaaaaata
tcatctggaattgtctctgaacatttttaaagacatttatgacaatgcctatgatgcaatt
gttcgtatggaaagacgttacaaaggaccgatagcttttgaatacactcacattaataat
aataaagaacgtgtgtggttaaaaagaagaattgaaacgccttataaagcaagtttaaac
gataatcaaaaaaagaacttttcaaaaaactcgcacacgtagaaggttttgaaaaatat
```
35
```
ttgcacaaaaattttgttggggctaaacgttttcaattgaaggcgtcgaTC
```

Sequence 2792
```
MLDLYDDYLQNPSSVPEDLQVLFSTIKTGEAHIEAKPTTDGGGSQAGDSTIKRVMRLIDN
IRQYGHLKADIYPVNPPERQNVPKLEIEDFDLDKETLEKISSGIVSEHFKDIYDNAYDAI
```
40
```
VRMERRYKGPIAFEYTHINNNKERVWLKRRIETPYKASLNDNQKKELFKKLAHVEGFEKY
LHKNFVGAKRFSIEGVDX
```

Sequence 2793
Contig_0440_pos_5821_6999
45 is similar to (with p-value 2.0e-24)
 >sp:sp|P23524|YHAD_ECOLI HYPOTHETICAL 42.1 KD PROTEIN IN RN
PB-SOHA INTERGENIC REGION (ORF 3) (F408). >pir:pir|JQ0614|JQ
0614 hypothetical 42K protein - Escherichia coli >gp:gp|D902
12|ECORNPBW_3 E.coli rnpB gene and ORFs. NID: g216630. >gp:g
50 p|U18997|ECOUW67_54 Escherichia coli K-12 chromosomal region
  from 67.4 to 76.0 minutes. NID: g606010. >gp:gp|AE000394|AE
000394_2 Escherichia coli K-12 MG1655 section 284 of 400 of
the complete genome. NID: g2367197.
```
atgtttaaaataattttttggaaaagagaaaaataaggtggttaagacaatgaaagtttta
```
55
```
gtagccatggatgaatttaatggaattatttctagttaccaagctaatagatatgttgaa
gaagcggtagcaagtcaaattgaagatgcagatatcgttcaagttccactatttaacggt
cgtcacgaattattagattcagtctttctttggcaatcaggaaataaatatcgtgtgagt
gcgcatgatgctgacatgaaagaaaccgaagcaatatatggacaaacggatagtggtatg
actattatcgaaggtcacttatttttaaatggcaaaaaacctattcaacatcgatcaagt
```

682

```
        tacggtttgggagaggttataaaagcagcattggacaatcatacagaacatcttgttatt
        tctttaggtggaataggaagttttgatggcggtgcaggcatgttgcaagcattgggtgca
        acattttatgatgatgaagcacaaattgtcgatatgaggaaaggtgcatatttaataaaa
        tatattagacgtattgatttatcaggtgttcatccacaattaacaaaggtaaacattcaa
5       ttaatgtcagatttctcaagtcgattgtatgggaaaaaagtgaaatcatgcaaacatac
        gaatcattagatttgtctcaaaatgaagcagccgagatagataatttaatttggtatttt
        agtgaattatttaagaatgaattgaaaatagcaatgggaccaatcgagcgcggtggtgct
        ggaggtggtatagcagctgtattaaatagtctctatcaagctgagattttaacaagccat
        gaattagtgaatcaaatcacacatttagaaaacttaattcaacaggcagatcttattatt
10      ttcggagaaggtttgaaagaagaagatcaaattctagagactacaacaatacgtatagca
        gaacttacccaacaatacagcaagccagctattgcaatttgtgctacaaatgataaattt
        gatttgtttgaatcattgaatgttacagcaatgtttaatacatttattgatatgcctgat
        tcatatacagattttaagatgggtattcaaatcagacattacacagtacaagcactaaaa
        ttattgaaaacgcaaataaacttaccgctttcatcctaa
15

        Sequence 2794
        MFKIIFGKEKNKVVKTMKVLVAMDEFNGIISSYQANRYVEEAVASQIEDADIVQVPLFNG
        RHELLDSVFLWQSGNKYRVSAHDADMKETEAIYGQTDSGMTIIEGHLFLNGKKPIQHRSS
        YGLGEVIKAALDNHTEHLVISLGGIGSFDGGAGMLQALGATFYDDEAQIVDMRKGAYLIK
20      YIRRIDLSGVHPQLTKVNIQLMSDFSSRLYGKKSEIMQTYESLDLSQNEAAEIDNLIWYF
        SELFKNELKIAMGPIERGGAGGGIAAVLNSLYQAEILTSHELVNQITHLENLIQQADLII
        FGEGLKEEDQILETTTIRIAELTQQYSKPAIAICATNDKFDLFESLNVTAMFNTFIDMPD
        SYTDFKMGIQIRHYTVQALKLLKTQINLPLSS*


25      Sequence 2795
        Contig_0440_pos_4831_4487
        is similar to (with p-value 1.0e-60)
         >pir:pir|I67760|I67760 transposase (insertion sequence IS10
        ) - Escherichia coli >gp:gp|S67119|S67119_2 BST=somatotropin
30      ...BST/beta-Gal fusion protein [Escherichia coli, LBB84, pla
        smid pXT107, IS10L/R-1, PlasmidTransposonInsertionMutant, 3
        genes, 1679 nt]. NID: g455674.
        atgcagattgaagaaaccttccgagacttgaaaagtcctgcctacggactaggcctacgc
        catagccgaacgagcagctcagagcgttttgatatcatgctgctaatcgccctgatgctt
35      caactaacatgttggcttgcgggcgttcatgctcagaaacaaggttgggacaagcacttc
        caggctaacacagtcagaaatcgaaacgtactctcaacagttcgcttaggcatggaagtt
        ttgccgccattctggctacacaataacaagggaagacttactcgtggctgcaaccctacta
        gctcaaaatttattcacacatggttacgctttggggaaattatga


40      Sequence 2796
        MQIEETFRDLKSPAYGLGLRHSRTSSSERFDIMLLIALMLQLTCWLAGVHAQKQGWDKHF
        QANTVRNRNVLSTVRLGMEVLRHSGYTITREDLLVAATLLAQNLFTHGYALGKL*


        Sequence 2797
45      Contig_0441_pos_2559_3467
        is similar to (with p-value 2.0e-39)
         >sp:sp|O54983|CRYM_MOUSE MU-CRYSTALLIN HOMOLOG. >gp:gp|AF03
        9391|AF039391_1 Mus musculus mu-crystallin (Crym) mRNA, comp
        lete cds. NID: g2745895.
50      atgaatgaagtcatttttagaagtagaaaaagctttgcaagctttttcagagaataagacg
        ataacgccattaagatatgttttgccatttaatgagcaaaatcgttatttagtgatgcca
        gcattatcagatgaattaaatatcgttggacttaaaacagtctcatttgcacctgaaat
        tcaaaaaaagggaaagcgactattactggatcagttattttaagtgactatgaaacagga
        gaaacattgtctatattagatggtggttttcttactaaagtaagaactggtgcaatttca
55      ggtgtagctactaaatatctagcaaaagaaacgctaaaacacttagtgtaataggggca
        ggtgtacaagctgaaggtttaattgaagcgatacttgctgttagagatattgaaaaatt
        cacatcgcaagtagaacgttcgaaaaagcagaaaaatttgctcaaaatatacgaaatcga
        tttaatattaaagtgagtgtatttagatcggcagatgaagcgatagacagtgcagatatt
        gtagttacagcaacaaatgcaaatcagcccgtttatactcattctttacatccaggcgtg
```

```
catttaaatgcagtcggatcctttaaaccagatatgcaagaaataccttcagaaacaatg
cttgttgctaataaaattgttgttgaatctatggaagcagctttagaagaaacaggtgat
ttaaaaattcctcaagcagaaggaatattaactaaaaatatgctacatagtgaattaggc
gacattatttctggtgaaaaagttggccgcgaaactgaagaagagtgacagtctttaaa
tcggtcggtctagcaattgtagatatcattgtggcacaatattttataaaaaattaata
caatcttag
```

Sequence 2798
```
MNEV LF VEKALQAFSENKTITPLRYVLPFNEQNRYLVMPALSDELNIVGLKTVS FAPEN
SKKGKATITGSVILSDYETGETLSILDGGFLTKVRTGAISGVATKYLAKENAKTLSVIGA
GVQAEGLIEAILAVRDIEKIHIASRTFEKAEKFAQNIRNRFNIKVSVFRSADEAIDSADI
VVTATNANQPVYTHSLHPGVHLNAVGSFKPDMQEIPSETMLVANKIVVESMEAALEETGD
LKIPQAEGILTKNMLHSELGDIISGEKVGRETEEEVTVFKSVGLAIVDIIVAQYFYKKLI
QS*
```

Sequence 2799
Contig_0441_pos_4023_4979
 unknown
```
atgaaaaaataataatcactggagcacttggtcaaattggtactgaattagttattaag
tgtagagaaagatatggaactgaaaatgtattagctactgatataagaaagccagaacca
cactctccagttaaaaacggtccttttgaaatcttagatgttacagatagaaatcgttta
tttgaaacagtgagatatttcaatgctgatacacttatgcatatggctgcactgctttct
gctacagcggaaaagaaaccactagttgcctgggatttgaatatgggggggacttataaat
acactcgaagcagcaagaaggtatcaattaaaatattttacgccaagttcaataggagct
tttggtatttcgacacctaaagtaaatacacctcaattaacaatacaacaaccaacgact
atgtatggtataaacaaagtaactggggaattattatgtcagtattattatgtcaaattt
ggagtagatacacgtagcgttagatttcctggtttaatttcacatgttaaagaacctggc
ggt ggca aacagattatgcagttgatatgtattttaaggcagttagaaaaggac tat
acaagtt tattaatcgttacacttatatggatatgatgtacatggaagatgcta tgat
gcgattatcaaattaatggaagaagacagtgttaaattaaaaacaagaaatggctataac
ttaagtgcgatgagtattgaacctgaaatgctaaaacaagctatacaagtatactatcca
gatttcaccttagattatgacattgatttagagcgtcaagatattgcattaagttggcca
gatagtatagatacgagctgtgctcaggaagagtggggatttgatcctaaatatgattta
ccaactatgactaaagtgatgttagaggcaattgagaaaaaacaaaaagaatgctga
```

Sequence 2800
```
MKKIIITGALGQIGTELVIKCRERYGTENVLATDIRKPEPHSPVKNGPFEILDVTDRNRL
FETVRYFNADTLMHMAALLSATAEKKPLVAWDLNMGGLINTLEAARRYQLKYFTPSSIGA
FGISTPKVNTPQLTIQQPTTMYGINKVTGELLCQYYVKFGVDTRSVRFPGLISHVKEPG
GGITDYAVDMYFKAVRKGHYTSYINRYTYMDMMYMEDAIDAIIKLMEEDSVKLKTRNGYN
LSAMSIEPEMLKQAIQVYYPDFTLDYDIDLERQDIALSWPDSIDTSCAQEEWGFDPKYDL
PTMTKVMLEAIEKKQKEC*
```

Sequence 2801
Contig_0442_pos_1526_2359
is similar to (with p-value 1.0e-60)
 >pir:pir|I67760|I67760 transposase (insertion sequence IS10
) - Escherichia coli >gp:gp|S67119|S67119_2 BST=somatotropin
...BST/beta-Gal fusion protein [Escherichia coli, LBB84, pla
smid pXT107, IS10L/R-1, PlasmidTransposonInsertionMutant, 3
genes, 1679 nt]. NID: g455674.
```
atgaagagtattttattttaagtaataatgtgaaaatattcacaaaaaaactaggagga
tttgctatgagtaaggaaatattcgatacttttaaatttaaatgtggtgccgaattaaaa
aatagagtattaatggcacccatgactatccaagctgggtattttgatggaagtgttaca
tcagaaatgattgattattatcaatttagagctggtgatgcttcagcaatcattgttgaa
agttgttttgttgaaaatcacggacgaggatttccgggagctataggtattgataatgat
gacaaaatacctggactcaaacgtttagcagaagcgattcaagctaagggatcaaaagcg
attttgcaactttatcatgccggaagaatggcaaatcctaaatttaatgaaggagagcag
ccgatatctgcgagcccccattgcagcattaagacctgatgctgtaccacctagagaaatg
```

```
acacatgctcaaatcaatcagatgattgatgactttggagaggctacacgtcgcgctata
gaagcggggtttgatggtgtcgaaattcatggcgccaacacatacttattacaacaattt
ttctctccacattctaatcggagacaagattcatggggaggcagtcgtgaaaaacgtaca
cgatttccaatcgaagttttgacaaaggttcaacacgtcgttgctgaaaaagaggcttct
cattttattataggatatcgattctcacctgaagaaattgaagaaccaggcatacgtttt
gaagataccatgtttttactaaatacattagcagaatgcttgactacaaaatag
```

5

Sequence 2802
MKSILFLSNNVKIFTKKLGGFAMSKEIFDTFKFKCGAELKNRVLMAPMTIQAGYFDGSVT
SEMIDYYQFRAGDASAIIVESCFVENHGRGFPGAIGIDNDDKIPGLKRLAEAIQAKGSKA
ILQLYHAGRMANPKFNEGEQPISASPIAALRPDAVPPREMTHAQINQMIDDFGEATRRAI
EAGFDGVEIHGANTYLLQQFFSPHSNRRQDSWGGSREKRTRFPIEVLTKVQHVVAEKEAS
HFIIGYRFSPEEIEEPGIRFEDTMFLLNTLAECLTTK*

10

15

Sequence 2803
Contig_0442_pos_13117_13461
is similar to (with p-value 5.0e-53)
   >sp:sp|P54524|YQIG_BACSU PROBABLE NADH-DEPENDENT FLAVIN OXI
DOREDUCTASE YQIG (EC 1.-.-.-). >gp:gp|D84432|BACJH642_230 Ba
cillus subtilis DNA, 283 Kb region containing skin element.
NID: g2627063. >gp:gp|Z99116|BSUB0013_132 Bacillus subtilis
complete genome (section 13 of 21): from 2395261 to 2613730.
   NID: g2634723.

```
atgcagattgaagaaaccttccgagacttgaaaagtcctgcctacggactaggcctacgc
catagccgaacgagcagctcagagcgttttgatatcatgctgctaatcgccctgatgctt
caactaacatgttggcttgcgggcgttcatgctcagaaacaaggttgggacaagcacttc
caggctaacacagtcagaaatcgaaacgtactctcaacagttcgcttaggcatggaagtt
ttgcggcattctggctacacaataacaagggaagacttactcgtggctgcaaccctacta
gctcaaaatttattcacacatggttacgctttggggaaattatga
```

20

25

30

Sequence 2804
MQIEETFRDLKSPAYGLGLRHSRTSSSERFDIMLLIALMLQLTCWLAGVHAQKQGWDKHF
QANTVRNRNVLSTVRLGMEVLRHSGYTITREDLLVAATLLAQNLFTHGYALGKL*

35

Sequence 2805
Contig_0445_pos_0_322
is similar to (with p-value 8.0e-17)
   >gp:gp|M89774|ECOLYSP_2 Escherichia coli lysine specific pe
rmease (lysP) gene, complete cds. NID: g466776. >gp:gp|AE000
305|AE000305_1 Escherichia coli K-12 MG1655 section 195 of 4
00 of the complete genome. NID: g1788479.

```
gtggcaactggaagtgtcatttctcaagctggcccaggaggagctatattagcttatata
ctaattggtattatgctttattttttaatgtcatcaataggagaattagcaactttctat
ccggtttctggttcttttagttcatactctaccagatttgttgattcgtcacttggtttt
acaatgggttggttgtattggggtatgtggtcacttgtaacaagtgtagatatcattgtt
gcttccaatgtattacaatattgggatgtatttaaagttttaaatccacttacatggagc
ttaattttcttaactctgttgt
```

40

45

Sequence 2806
VATGSVISQAGPGGAILAYILIGIMLYFLMSSIGELATFYPVSGSFSSYSTRFVDSSLGF
TMGWLYWGMWSLVTSVDIIVASNVLQYWDVFKVLNPLTWSLIFLTLLX

50

Sequence 2807
Contig_0447_pos_14098_15573
is similar to (with p-value 6.0e-40)
   >gp:gp|X93084|MBFMDSUBS_1 M.barkeri fmdE, fmdF, fmdA, fmdC,
   fmdD, fmdB, orf4, orf3, orf2, and orf1 genes. NID: g1124956

55

```
atgaaaacattgattactcttattgtcatgatatcttttgttttagcgagctgtgcgggc
```

```
acatcaagtacagacaaagacaccctcaatgttgaaatacctttgaaaactaaatcaatt
gcaccttacgaaactgatatcccagttaaaacaggtgccttggaatcgcttttttaaaatg
tcgaagaatggtaaagtaaaacctttattagtcaaaaattatcatcaagtatctgacaat
caactagaactcactttaaaagataatattaaatttcaaaacggtcatcatttaacaggc
gaagctgtaaaacgtagtctcgaagaaggaatgaaaaaaagtgatttgttaaaaggatca
cttcctattaaatcaatcaatgctcatggacaaaaagtcacaatcactactaaagaacct
tatccagaattaatgtctgaactcgcaagcccatttgctgctatttacgacacaaaagct
aaaaacaaagtaactgatcaacctgttggtacgggtccttataaaattgatcagtataaa
cgttcgcaaaaaatcgtactaaaacaattcaaagactactggcaaggtacgccaaaatta
aaaagaattaatgtcacttatcatgaagatggtaatactcgtgttgatcacttattatca
ggcaaatcagatttgactactgatgttccaattgaacgcgttgatgatgtaaaaaaatct
aacaaagcaaacattcaaagtacatcaggctttagaacgcatttaatgttatacaatcat
gatagtaaaaaagttaataaaaaagtaagagaagcactagatatgattattaatcgaaaa
gacattgctaaaaatgtttctaaaaattatgctgagccagcatcaggtcctttaaccat
cgattaaaatcattagaaaaagaggaaattcaatcacaagacatcaagcgtgcaaaagaa
ctttagctcaagaaggttattccaaatcgcatcctcttaaattaaacatggtcacatac
gatggcagaccagaattgcctaaaattggacaggtgatacaatctgaagctaaaaaggca
aatgttgatatacaattacgcaatgtagatgatatcgaaggatatctcaaaaacaaacag
agttgggatgtttcaatgtatagttatttaagtgtgccacgtggtgatacaggttatttc
tttaacactgcatacttacctgatggagcattaaataaaggtaattatagtaatactaaa
gtcactcagttaattaaagaattaaatactactttcggtgacaagcaacgtggtcaagtg
actaatgaaatactaaatgaatctaaaaaagatattcctaacagctatatcacatacaac
tctcaaatagatggtgtgaataataaagtaagacattttaatgttacaccagaatctatc
tatttaattgattataaattaagtaaaaaagaataa
```

Sequence 2808
```
MKKLITLIVMISFVLASCGGTSSTDKDTLNVEIPLKTKSIAPYETDIPVKTGALESLFKM
SKNGKVKPLLVKNYHQVSDNQLELTLKDNIKFQNGHHLTGEAVKRSLEEGMKKSDLLKGS
LPIKSINAHGQKVTITTKEPYPELMSELASPFAAIYDTKAKNKVTDQPVGTGPYKIDQYK
RSQKIVLKQFKDYWQGTPKLKRINVTYHEDGNTRVDHLLSGKSDLTTDVPIERVDDVKKS
NKANIQSTSGFRTHLMLYNHDSKKVNKKVREALDMIINRKDIAKNVSKNYAEPASGPFNH
RLKSLEKEEIQSQDIKRAKELLAQEGYSKSHPLKLNMVTYDGRPELPKIGQVIQSEAKKA
NVDIQLRNVDDIEGYLKNKQSWDVSMYSYLSVPRGDTGYFFNTAYLPDGALNKGNYSNTK
VTQLIKELNTTFGDKQRGQVTNEILNESKKDIPNSYITYNSQIDGVNNKVRHFNVTPESI
YLIDYKLSKKE*
```

Sequence 2809
Contig_0447_pos_20772_21911
>gp:gp|Z82038|CTZ82038_4 C.thermosaccharolyticum etfB, etfA
, hbd, thlA and actA genes. NID: g1667352. >gp:gp|Z92974|TTB
CSOPRN_6 T.thermosaccharolyticum BCS operon DNA. NID: g19033
26.
```
gtgtttggtggtgtatttaaggatatacctgcctatgaactaggtgcaacagttattcgt
caaattttagaacatagtcaaatagatcctaatgaaatcaatgaagttattctaggaaac
gtattacaggcaggtcaaggacaaaatcctgctcgtattgctgcgattcatggtggtgtg
ccagaagcggtaccttcttttactgtaaataaagtttgcggttctggattaaaagcgatt
caacttgcctatcaatctattgtagcgggagataatgagattgttatcgctggaggcatg
gaaagtatgtctcaatctccaatgcttcttaaaaatagtcgtttcggttttaaaatggga
aatcaaactttagaagatagtatgatagctgatggtttaactgataagtttaatgattac
catatgggtatcacagccgaaatctagttgaacagtatcagattagtcgtaaagaacaa
gatcaattgcattcgattctcaacaaaaagcatcacgtgcacaacaagctggtgtattt
gatgctgaaattgtacctgtagaggtaccacaacgtaaaggcgacccctaattatttct
caagatgaaggcattagacctcaaacgacaattgataagttagcacaactccgtccagca
tttaaaaaagatggatcagtaactgctggtaatgcatccggtatcaatgacggtgctgct
gctatgctcgttatgacggaggacaaagcgaaagcattgggcttacaacctatagctgta
ttagatagttttggtgcgagtggtgtggcgccttcaattatgggtattggtccagttgaa
gcgatacataaagctttaaaacgttctaataaagtgataaatgatgttgatattttttgaa
ttaaacgaagctttgcagcgcaatcaattgctgtaaaccgtgagttgcaattaccgcaa
gataaagtcaatgttaatggtggtgcgattgcactaggacatccgatagggggcttcgggt
```
```

gcgcgtactttagtttcattattacatcaattaagtgatgctaagccaacaggtgtggca
tctttatgtatcggtggcggtcaaggtatcgctacggttgtatctaaatatgaagtttaa

Sequence 2810

VFGGVFKDIPAYELGATVIRQILEHSQIDPNEINEVILGNVLQAGQGQNPARIAAIHGGV
PEAVPSFTVNKVCGSGLKAIQLAYQSIVAGDNEIVIAGGMESMSQSPMLLKNSRFGFKMG
NQTLEDSMIADGLTDKFNDYHMGITAENLVEQYQISRKEQDQFAFDSQQKASRAQQAGVF
DAEIVPVEVPQRKGDPLIISQDEGIRPQTTIDKLAQLRPAFKKDGSVTAGNASGINDGAA
AMLVMTEDKAKALGLQPIAVLDSFGASGVAPSIMGIGPVEAIHKALKRSNKVINDVDIFE
LNEAFAAQSIAVNRELQLPQDKVNVNGGAIALGHPIGASGARTLVSLLHQLSDAKPTGVA
SLCIGGGQGIATVVSKYEV*

Sequence 2811
Contig_0447_pos_24387_22894
 >pir:pir|S57636|S57636 5-methyltetrahydropteroyltriglutamat
e--homocysteine S-methyltransferase (EC 2.1.1.14) - Madagasc
ar periwinkle >gp:gp|X83499|CRMETS_1 C.roseus MetE mRNA for
methionine synthase. NID: g886470.
gtgggcggtttaggattagatttggtacacgataacggctataacttaaaacaaattgaa
gatggtaatttcgatcaaagtaaagcactttatgcaggaatcattgatggcagaaatgta
tgggcggctgatattgaagcaaaaaaacaattaatcgaaacattacaacaacacacacaa
cagttagtcattcaaccttcatcatcacttctacatgttccagtatcactagatgatgaa
acacttgatgaatctattgcagaaggattaagttttgctactgaaaagttagacgaatta
gatgcattgcgtcgcttattcaatgataatgacctgtcaaaatatgaacattacaaagca
cgttatgaacgcttccaaagccaatcatttaaaaatcttgaatacgattttgaaagtgta
ccgacacatcgtaaatcaccatttgctaaacgcaagcaattacaaaaccaacgtctaaat
ttaccagatttacctacgactacgattggttcattccctcaaactcgtgaagtccgtaaa
tttagagcagattggaaaaataatcgcattacagatgctgagtatcaagaatttttacaa
aatgaaatcgctcgttggattaaaatccaagaagatattggtttagacgtacttgtacac
ggtgagtttgagcgtaacgacatggttgaattttttggggaaaaacttcaaggtttccta
gtaacaaaattcggttgggtacagtcttacggttctcgtgcagttaaaccaccagttata
tatggagatgtgaagtggactgccccacttactgtaaaagaaacagtatacgcgcaaagc
ttaactgacaaaccagtaaaaggtatgcttacaggtccagttactattttaaactggtca
tttgaacgcgtagacgtacctagaaaggttgttcaagatcaaattgccttagccattgat
gaagaagtgttagctttagaagaagcaggtatcaaagtcattcaagttgacgaaccaacg
ttgcgtgaaggcttacctctacgctctgaatatcacgaacaatatttagaagatgctgta
cattcatttaaattagcgacatcatctgtgcatgatgaaacacaaattcatacacatatg
tgctattctcaatttggacaaatcattcatgctatccatgatttagatgcagatgttatt
tctattgaaacttctcgtagtcacggagacttaattcaagatttcgaagatattaattat
gatttaggaattggattaggtgtttacgatattcacagtccacgtattccaaccgaagaa
gagattactactgcaattaatcgttcattacaacaaattgatcgttctctattctgggtt
aacccagactgtggtctaaaaacacgtaaagaaaatgaagtgaaagatgcattaacagtg
ttagtcaatgcagttaaaaagaaacgtcaagaatctgaatcaacaacagcataa

Sequence 2812
VGGLGLDLVHDNGYNLKQIEDGNFDQSKALYAGIIDGRNVWAADIEAKKQLIETLQQHTQ
QLVIQPSSSLLHVPVSLDDETLDESIAEGLSFATEKLDELDALRRLFNDNDLSKYEHYKA
RYERFQSQSFKNLEYDFESVPTHRKSPFAKRKQLQNQRLNLPDLPTTTIGSFPQTREVRK
FRADWKNNRITDAEYQEFLQNEIARWIKIQEDIGLDVLVHGEFERNDMVEFFGEKLQGFL
VTKFGWVQSYGSRAVKPPVIYGDVKWTAPLTVKETVYAQSLTDKPVKGMLTGPVTILNWS
FERVDVPRKVVQDQIALAIDEEVLALEEAGIKVIQVDEPTLREGLPLRSEYHEQYLEDAV
HSFKLATSSVHDETQIHTHMCYSQFGQIIHAIHDLDADVISIETSRSHGDLIQDFEDINY
DLGIGLGVYDIHSPRIPTEEEITTAINRSLQQIDRSLFWVNPDCGLKTRKENEVKDALTV
LVNAVKKKRQESESTTA*

Sequence 2813
Contig_0450_pos_7632_8510
 No hits found
atgaaagcgccagttctggtatcaggtactgacggtgtgggtacaaagttaaaattagca

```
      attgactatggaaagcatgacacaattggtattgatgctgtcgcaatgtgtgtaaatgat
      attttaacaacaggtgctgaacctttatactttttagactatattgccacgaataaagta
      gtgccaagtactatagagcaaatcgttaaaggtataagtgacggttgcgaacaaaccaat
      acggcacttataggcggtgaaactgctgaaatgggagaaatgtatcatgaaggtgaatat
 5    gatattgctggttttgcagtaggagcggtagagaaagaggactatattgatggttcaaat
      gttgaagaaggacaagcaattattggtttagcttcaagtggtattcattcaaatggctat
      agtctagttagaaaaatgataaaagaatcaggagttcaattacatgatcaatttaatggt
      caaacctttttagaaaccttccttgcaccaacaaaattgtatgtaaagcctattcttgaa
      ttaaagaaacatattgatatcaaagcgatgagccatattactggtggaggtttctatgaa
 10   aatattccgcgtgccccttcctaaaggtttatcagcaaaaatagatacacaatcattccca
      acgttggaagtctttaattggcttcaaaaacagggcaacatttcaacgaatgaaatgtat
      aacatatttaatatgggtattggatatacaattattgttgacaaaaaagatgttcaaaca
      acattaacaacgttacgtgcaatggatacaactgcatatgaaattggtgagattataaaa
      gatgatgatacacctattcatttattggaggtagaatag                    .·
 15
      Sequence 2814
      MKAPVLVSGTDGVGTKLKLAIDYGKHDTIGIDAVAMCVNDILTTGAEPLYFLDYIATNKV
      VPSTIEQIVKGISDGCEQTNTALIGGETAEMGEMYHEGEYDIAGFAVGAVEKEDYIDGSN
      VEEGQAIIGLASSGIHSNGYSLVRKMIKESGVQLHDQFNGQTFLETFLAPTKLYVKPILE
 20   LKKHIDIKAMSHITGGGFYENIPRALPKGLSAKIDTQSFPTLEVFNWLQKQGNISTNEMY
      NIFNMGIGYTIIVDKKDVQTTLTTLRAMDTTAYEIGEIIKDDDTPIHLLEVE*

      Sequence 2815
      Contig_0450_pos_13563_14528
 25   is similar to (with p-value 9.0e-27)
        >gp:gp|Z99108|BSUB0005_100 Bacillus subtilis complete genom
        e (section 5 of 21): from 802821 to 1011250. NID: g2633055.
        >gp:gp|D78508|D78508_5 Bacillus subtilis DNA for YfiO, YfiP,
        YfiN, YfiM, YfiL, YfiK, YfiJ, YfiI, YfiH, complete cds. NID
 30     : g1817531.
      gtgggtaaaggcaacaatcttggcaaaaatgctgctctcaaaacttataatgagaatcaa
      gcgcaacatctgcttaaacagaatcaattacaaggatactttgtcttcgaccgtggtatg
      accgatacttttttataaagatggtagccttcctataactatttatacatatgatgaacaa
      tcaagtaacagtgtcgtggttaatcaattaacacgctcagtttatgaccgtttaatgtta
 35   tcaatgggcggtgtgctaagctttaatcaattagctaaagatccttcaaatgaagacgta
      gcaatgacattgattgatatgttatttaccggtttaaatcgttcaggttcatttaatttt
      gaacccatacatatttatgacaccagtagttattatgtagtcactggatttctttttgtct
      atctttatattgtgtttatcactttatacagtactaaaaatgaatcaagaaactgcactc
      aaagaacgcttgcaaatgtttcatttttcctttgaaaagctcacgatagttcgaggtatc
 40   attgcatggtttttattcactcatatgggcatttattggctttatctggattactcatgct
      ctaaatgccccatttgaaaaatacaattggccaacggtagctttacaactcacttattac
      gttacgttactcgtcctatgcttattacttatagacttaattactcgttcatggataaac
      tttctactcaaattattacttagctttgttatcgttatttttctgggataattatccct
      actatcttctttaaacacatgcttaatgatgtaatcattacacaaccatttagtttggtt
 45   actaatcaaatgttagaaataacactcaataactatatttagacacacatccagcattt
      tatttaagttttattacactattgatactattcatcattgttttagtatggaggtatcgc
      cgatga

      Sequence 2816
 50   VGKGNNLGKNAALKTYNENQAQHLLKQNQLQGYFVFDRGMTDTFYKDGSLPITIYTYDEQ
      SSNSVVVNQLTRSVYDRLMLSMGGVLSFNQLAKDPSNEDVAMTLIDMLFTGLNRSGSFNF
      EPIHIYDTSSYYVVTGFLLSIFILCLSLYTVLKMNQETALKERLQMFHFSFEKLTIVRGI
      IAWFYSLIWAFIGFIWITHALNAPFEKYNWPTVALQLTYYVTLLVLCLLLIDLITFSWIN
      FLLKLLISFVIVIFSGIIIPTIFFKHMLNDVIITQPFSLVTNQMLEITLNNYILDIHPAF
 55   YLSFITLLILFIIVLVWRYRR*

      Sequence 2817
      Contig_0450_pos_15460_16149
      is similar to (with p-value 1.0e-18)
```

688

>sp:sp|P22082|SNF2_YEAST TRANSCRIPTION REGULATORY PROTEIN S
NF2 (SWI/SNF COMPLEX COMPONENT SNF2) (REGULATORY PROTEIN SWI
2) (REGULATORY PROTEIN GAM1) (TRANSCRIPTION FACTOR TYE3). >p
ir:pir|S15047|S15047 SNF2 protein - yeast (Saccharomyces cer
evisiae) >gp:gp|X57837|SCSNF2_1 Yeast GAM1/SNF2 gene for a n
uclear protein required for transcription of STA1 gene. NID:
g4499. >gp:gp|X89633|SCVCOSMGN_20 S.cerevisiae DNA for VPH1
, MOD5, CAP20, ORF1 and SNF2 genes. NID: g1279694. >gp:gp|Z7
5198|SCYOR290C_1 S.cerevisiae chromosome XV reading frame OR
F YOR290c. NID: g1420643. >gp:gp|D90459|YSCRIC1_1 Yeast RIC1
gene (regulatory gene for phospholipid synthesis), complete
cds. NID: g806531. >gp:gp|M61703|YSCSNF2A_1 Saccharomyces c
erevisiae SNF2 protein gene, complete cds. NID: g172631.
atgatagatattcaaaatgtttccaaaagctataaaaagaagcatattttcgattcccta
gatatgcaatttcaaaatcataaaattactattttattaggtgaaaatggtgctggaaaa
tctacattattgcgtttaattgcaggtattgagaatgcagacgaaggacgtattcaatac
ttcaatcaatatttgtcaagacgtcgaatacgtcatattgtaggctatgtccctcaagac
atcgcactattcgagcatatgactgtcatggagaacattgagttttttcaagtcactttgt
gaaaatcctatttcagatgaaacacttcattcttatttatcacaattaaattttactgat
acaaaagtgaaagtatctaacctttctgggggaaataaacgtaaagtcaatattatgata
ggtctacttagtcggcctaaaatacttattctagatgagccaacagaaggcattgattta
gaatcaagatatgatattcacaacttattacaacaaatgaccgatcaatgtttaatcatc
atgacgacacatcatttagacgaagttgaagcactagcagatgatattaaagttataggt
caaaatccttttttatcatgatatttagaaaataaaggttggtcttttaaaaaatatgca
aatgccttagctgataatacgaaatcttaa

Sequence 2818
MIDIQNVSKSYKKKHIFDSLDMQFQNHKITILLGENGAGKSTLLRLIAGIENADEGRIQY
FNQYLSRRRIRHIVGYVPQDIALFEHMTVMENIEFFKSLCENPISDETLHSYLSQLNFTD
TKVKVSNLSGGNKRKVNIMIGLLSRPKILILDEPTEGIDLESRYDIHNLLQQMTDQCLII
MTTHHLDEVEALADDIKVIGQNPFYHDILENKGWSFKKYANALADNTKS*

Sequence 2819
Contig_0450_pos_20344_21786
is similar to (with p-value 5.0e-85)
>sp:sp|P43848|PUR5_HAEIN PHOSPHORIBOSYLFORMYLGLYCINAMIDINE
CYCLO-LIGASE (EC 6.3.3.1) (AIRS) (PHOSPHORIBOSYL-AMINOIMIDAZ
OLE SYNTHETASE) (AIR SYNTHASE). >pir:pir|G64122|G64122 5'-ph
osphoribosyl-5-aminoimidazole synthetase (purM) homolog - Ha
emophilus influenzae (strain Rd KW20) >gp:gp|U32822|U32822_2
Haemophilus influenzae Rd section 137 of 163 of the complet
e genome. NID: g1574265.
atgaaaagagcagccaatttttctgtacctggagctggtaaaacggctatgatgtatggc
acatttgctttttttgtctagtgaaataaagcggaaagttgataaattaatagttattttct
ccaattaatgcatttgaagcttggcgttcagaatttattgaagttttttcaagataaaaga
gaattacacttttatgaacctaagagataaaaaatataatgatttaggtaaagtacgaaca
gattggggaagtgcaaatgtcattgtttttgaattttgaagcaatacaaaagtatgtaggg
gttttaaatgaacttattaatgataagacaatgatagtttatgatgaggttcataggata
aaaggtattaatagtagcagagcaagttatgcattaactttaggtcctaaaagttattac
agatacgttttaactggtaccccaattccaaatagttatcaagacatatttaacttctta
aatattttatataaagatgagtatgatacttattttggttggaatgttgctgatttacaa
aatcctgatcctaatgaaattaatgacaagttgaacccttttttctggcgtacaaataag
aacgatttggaagtgcctcaagcagaaatgatattatattatgtgttaagcctagtaat
attcaaattgaattagcaaaagcgatatacgaaatgaatctgggatactagcgatttat
ataaggttactacaagcttcaactaacccagaattattgcaaaagaatattaattatagc
gaactaggaatgttgaatgatgaattaaatttcgatttggataaagcattaaataaagaa
gaagaaaatcaagtaaaacaacaaatttataattctttttgatttaaaaaaatgtaacttct
ccaaaattcgaaaaaggtattgaattgattgaaaacttagttagccaaggaaaaaaagtt
ctagtatggggattgtttgtaggtacaatgaataaaatcaataagaggttactagaaagt

gatattaattcaatattgatttatggagaaacacctaaagaagatagggtagatatgatc
aataattttaggaatggaaatgcacaagttctaatatctaatcctaatacattaggcgag
tccatatctttacatcagacagtacatgatgcaatatattttgaatataactttaattta
acgtttatgttgcaatcacgtgatagaatacatcgtttagggttaaataataatcaatat
acaaggtattattatttgatgtctgaaggggatagagcccataaaggtttttatcgataaa
gcagtttataacagactgaaggaaaaagaagatgtaatgttaaatgctattgatggaaat
actttaaagccaatgattgaagatgattacttagaagatgttaagaaaattattattgaa
tga

Sequence 2820
MKRAANFSVPGAGKTAMMYGTFAFLSSEIKRKVDKLIVISPINAFEAWRSEFIEVFQDKR
ELHFMNLRDKKYNDLGKVRTDWGSANVIVLNFEAIQKYVGVLNELINDKTMIVYDEVHRI
KGINSSRASYALTLGPKSYYRYVLTGTPIPNSYQDIFNFLNILYKDEYDTYFGWNVADLQ
NPDPNEINDKLNPFFWRTNKNDLEVPQAENDIILCVKPSNIQIELAKAIYENESGILAIY
IRLLQASTNPELLQKNINYSELGMLNDELNFDLDKALNKEEENQVKQQIYNSFDLKNVTS
PKFEKGIELIENLVSQGKKVLVWGLFVGTMNKINKRLLESDINSILIYGETPKEDRVDMI
NNFRNGNAQVLISNPNTLGESISLHQTVHDAIYFEYNFNLTFMLQSRDRIHRLGLFNNQY
TRYYYLMSEGDRAHKGFIDKAVYNRLKEKEDVMLNAIDGNTLKPMIEDDYLEDVKKIIIE
*

Sequence 2821
Contig_0451_pos_806_1621
is similar to (with p-value 1.0e-20)
 >gp:gp|U67964|EVU67964_2 Ectromelia virus H14-B and H14-E g
enes, complete cds. NID: g2145123.
atgcattatataaaatttattgagtcaaaagataatacaaaactttatatgaaagtgaat
gatattcaagatgcaaaagcgaatatcattatagctcatggtgtggcagaacatttagat
cgttatgatgagataacagcatatttaaatgaagcgggtttttagtgttattagatatgat
caaagagggcatggtcgttctgaaggcaagcgtgccttttatagcaatagtaatgaaatt
gtcgaagatttagatgcgataataaaattatgtgaagtcaaactttgaaggtaaagtttac
ttaatcggtcatagtatgggtggttatacagtcactttatatggaacgaaacatccaaat
acagtgaatggtattataacttctggagcattaacacgttataataataaactatttggc
aatcctgatagaaacatatcacctgatacttatatagaaacaatttaagtgaggggta
tgttctgatttagaggtaatggaaaaatataaacttgatgatttgaatgcgaaacaaatc
tctatggggctcgtctttttcaataatggatggtgttaggtatttgaaagacaatgctcaa
caatttacagataatattttgatattgcatggcaaggaagatgggctagtaagctatgta
gattcttttacagctttatcaagaaataggatcagcacataaatcattacacatctstgat
cgtttggagcatgaaatatttaatgaaagttcttataatagaactattttttaacgaagtt
attgaatggcttgaaacggaattaacttataactaa

Sequence 2822
MHYIKFIESKDNTKLYMKVNDIQDAKANIIIAHGVAEHLDRYDEITAYLNEAGFSVIRYD
QRGHGRSEGKRAFYSNSNEIVEDLDAIINYVKSNFEGKVYLIGHSMGGYTVTLYGTKHPN
TVNGIITSGALTRYNNKLFGNPDRNISPDTYIENNLSEGVCSDLEVMEKYKLDDLNAKQI
SMGLVFSIMDGVRYLKDNAQQFTDNILILHGKEDGLVSYVDSLQLYQEIGSAHKSLHIYD
RLEHEIFNESSYNRTIFNEVIEWLETELTYN*

Sequence 2823
Contig_0455_pos_4830_4159
is similar to (with p-value 4.0e-17)
 >sp:sp|Q10092|YAOD_SCHPO HYPOTHETICAL 24.2 KD PROTEIN C11D3
.13 IN CHROMOSOME I. >gp:gp|Z68166|SPAC11D3_13 S.pombe chrom
osome I cosmid c11D3. NID: g1107889.
gtgtatatattcatgagtaaaaaagtttatttgttttaacaagtacaagtcaatttaca
gacggtacagaaactggattatggttagaagaagctggagcaccatataatatattgact
gaugaaggtatcaatgttgatgttatttctattaaaggtggaaaagtaaatcttgatcct
aattctgtttctaatgaatcactgaatcagtatgctaaattcgtgtcacacttaaacgat
acacctagtatcgaaaatgtaaatgcagatgagtatgacgctatttatctaccaggtgga
catggtactgtatacgattttgccaataatgagaaattagctgatattttacttcaattt

```
       aaaaatagtaataaaataatctcttcagtatgtcatggacctagtgcgtttgtaggtgta
       aaagatgcaaataatcactatctagtagatggtgtcaaaataacttcatttactgatagt
       gaagaaaaagcaatgggatttgaaaataaagtaccatttttaactcaatctaaattagaa
       gagcaaggtgcaaattttgtagtgaaagatgactttacatctcacgtagaaaaagacggt
5      caatttatcactggacaaaatccccaatcaagtgaagacattggtaaagcacttgcaaat
       gaattaaaataa
```

Sequence 2824
VYIFHSKKVLFVLTSTSQFTDGTETGLWLEEAGAPYNILTEEGINVDVISIKGGK.NLDP
10   NSVSNESLNQYAKFVSHLNDTPSIENVNADEYDAIYLPGGHGTVYDFANNEKLADILLQF
       KNSNKIISSVCHGPSAFVGVKDANNHYLVDGVKITSFTDSEEKAMGFENKVPFLTQSKLE
       EQGANFVVKDDFTSHVEKDGQFITGQNPQSSEDIGKALANELK*

Sequence 2825
15   Contig_0458_pos_1499_1038
       is similar to (with p-value 4.0e-51)
         >sp:sp|O31408|ARGR_BACST ARGININE REPRESSOR. >gp:gp|Y09546|
       BSARGR_1 B.stearothermophilus argR gene. NID: g2369705.
       gtgttaattgtgccaaaaaagtcagtgagacatataaaaataagagagataatttcaaat
20   gaacaaatagaaacacaagatgaactagttaaacgtttgaatgagtatgatttaaatgtt
       acacaagctactgtttcacgagatattaaagaattgcaattaattaaagttcctgcacct
       acagggcaatatgtttatagtttaccaaatgatcgtagatatcatccattagagaagttg
       ggtagatatttaatggattcatttgtaaacattgagggtactggtaatctactagttctt
       aaaacgcttcctggtaatgctcaatccattggtgctatacttgatcaaattgattgggat
25   gaggtacttggtacaatttgtggtgatgatacatgcttacttatttgtcgagacgaagaa
       gcgagtgaagaaatcaaaactcgaattttcaatttattataa

Sequence 2826
VLIVPKKSVRHIKIREIISNEQIETQDELVKRLNEYDLNVTQATVSRDIKELQLI.VPAP
30   TGQYVYSLPNDRRYHPLEKLGRYLMDSFVNIEGTGNLLVLKTLPGNAQSIGAILDQIDWD
       EVLGTICGDDTCLLICRDEEASEEIKTRIFNLL*

Sequence 2827
       Contig_0460_pos_2481_3032
35   is similar to (with p-value 2.0e-28)
         >sp:sp|P26646|YHDH_ECOLI HYPOTHETICAL 34.7 KD PROTEIN IN MR
       EB-ACCB INTERGENIC REGION (ORF1) (O324). >pir:pir|JS0688|JSO
       688 hypothetical 35K protein (fabE 5' region) - Escherichia
       coli >gp:gp|M80458|ECOACOAC_1 E.coli biotin carboxylase and
40   biotin carboxyl carrier protein (fabE) and ORF1 35 kDa prote
       in genes, complete cds. NID: g145172. >gp:gp|U18997|ECOUW67_
       183 Escherichia coli K-12 chromosomal region from 67.4 to 76
       .0 minutes. NID: g606010. >gp:gp|AE000404|AE000404_7 Escheri
       chia coli K-12 MG1655 section 294 of 400 of the complete gen
45   ome. NID: g2367207.
       atgtctattgaaggtaaagaagtgcttgtacgaggtgccactggaggcgtcggaacgatt
       tcattactcatgttaaataacttagggtatgatgttattgcaagtacgggtagagatgac
       gccggagaaaaacttaaaaagcttggtgctaaagaagtaattggccgtttaccagaagat
       aatagtaaaccattagagaagagaacatggcaggcagccattgatccagttggtggtgaa
50   aacttaccgtacatcgtcaaacgattggataacaacggaagtgttgcattaattggcatg
       actggtggtaacaattttgaaacaaccgtctttcctttcattttaagaggagcaagtata
       attggtatcgattcagtatttactccaattaaactaagaaaacgtgtttggagaagactt
       gcaaaagacttaaaaccacaacaattacatgacatcaaacatgttatttcattcgatgaa
       atcccaaaagccatcgatcaagtcatcaatcataataatactggacgtattgtcattgat
55   ttcaatgtttaa

Sequence 2828
       MSIEGKEVLVRGATGGVGTISLLMLNNLGYDVIASTGRDDAEEKLKKLGAKEVIGRLPED
       NSKPLEKRTWQAAIDPVGGENLPYIVKRLDNNGSVALIGMTGGNNFETTVFPFILRGASI
```

IGIDSVFTPIKLRKRVWRRLAKDLKPQQLHDIKHVISFDEIPKAIDQVINHNNTGRIVID
FNV*

Sequence 2829
Contig_0460_pos_2826_2440
is similar to (with p-value 5.0e-19)
>sp:sp|P26646|YHDH_ECOLI HYPOTHETICAL 34.7 KD PROTEIN IN MR
EB-ACCB INTERGENIC REGION (ORF1) (O324). >pir:pir|JS0688|JS0
688 hypothetical 35K protein (fabE 5' region) - Escherichia
coli >gp:gp|M80458|ECOACOAC_1 E.coli biotin carboxylase and
biotin carboxyl carrier protein (fabE) and ORF1 35 kDa prote
in genes, complete cds. NID: g145172. >gp:gp|U18997|ECOUW67_
183 Escherichia coli K-12 chromosomal region from 67.4 to 76
.0 minutes. NID: g606010. >gp:gp|AE000404|AE000404_7 Escheri
chia coli K-12 MG1655 section 294 of 400 of the complete gen
ome. NID: g2367207.
atgaaaggaaagacggttgtttcaaaattgttaccaccagtcatgccaattaatgcaaca
cttccgttgttatccaatcgtttgacgatgtacggtaagttttcaccaccaactggatca
atggctgcctgccatgttctcttctctaatggtttactattatcttctggtaaacggcca
attacttctttagcaccaagcttttttaagttttttcttcggcgtcatctctacccgtactt
gcaataacatcataccctaagttatttaacatgagtaatgaaatcgttccgacgcctcca
gtggcacctcgtacaagcacttctttaccttcaatagacatacctgatttttcaagttgt
tctatagctaaacctgctgtaaaatag

Sequence 2830
MKGKTVVSKLLPPVMPINATLPLLSNRLTMYGKFSPPTGSMAACHVLFSNGLLLSSGKRP
ITSLAPSFLSFSSASSLPVLAITSYPKLFNMSNEIVPTPPVAPRTSTSLPSIDIPDFSSC
SIAKPAVK*

Sequence 2831
Contig_0463_pos_814_2568
is similar to (with p-value 7.0e-69)
>gp:gp|Z47210|SPDEXCAP_7 S.pneumoniae dexB, cap3A, cap3B an
d cap3C genes and orfs. NID: g1658316.
atgataataagttatttacacaatataaataaattgaaaaactctaattttttagacttg
actcaaaataaaaatggaaacatgatatactttaaaaaagaggtaatactaaaaatgata
gataattggatagatgtattagatgaaagtttagtcaaagattttttataataatcaaact
tccgaagagcaacaagaaggacttgatactacactgtcttttggcacggctggtattaga
gggaaattcggtttaggcgaaggccgattaaataagttcaccgtatctaaagtagcgtta
ggctttgcccattatttaacatcaagtatcgcgcatcctgtcgtcgtcatacattatgac
acaagacacttatcacctgattttgctcaaattatcgctaatattctagcaagtcttgat
attaaagtttatcttgctgatacatacagaacaacacctgatttatcatttgcagtcaga
tacttacaggcggatgcgggtattatgattacagctagccataatcctaaagattataat
ggaatcaaagtgtatggagaagatggtgctcaattatcaaccgacgattccgcacgacta
agcacatatatcgataagttaggtcatccgcttcatattaatttacctagtttaactact
gaacaacaaacattaattcattcagtcccgagcgaagttagagaagattatttcaaaaac
gtacaagacttagttggaactattccacagtctgatttgaaagttgtctttacaagcttg
catggtacgagtgtgccagttgtacctgacatcttatcttctcttaactttaatcaattt
gagttagttgcatcacaatgtgaacctgattcagatttcagctctgtagtcagtgcaaat
ccagaggatcataaagcgtttgatcaatcgatagaacttgctaatcatattgatgctgat
ttacttattggcacagatcccgatgcagaccgtttaggaatagttgaacgtgatgctgaa
ggtaacatccactattacaacggaaatcagattggtgcacttttgttaaattatcgtatc
aaacaaacagaaggattacctaacagaattatgttccaatcaattgtaagtggtggctta
gctaaatctcttgctcaatatcataatgtcaatttcaaagaagtttttgacaggttttaaa
tatatcgcagctgaaataagacatctgtctcctgaacaaaactttattttttggctacgag
gaaagttatggcttcttagcccgtcctttcgtgagagataaagatgcgattcaaattgtg
ccattaatgattaagtatgcagctgaattaaaaaacaaaggacgcatgcttaaagatgaa
ttagaagacattactcgaaatgtcggtaactttaatgacaagcttttctcacatacattt
gaaggtactcaaggcaaggcaaaaatcgaaaatattatgactcaattcagaagtgaaaca

ccttcggaaatgtgtggtcttaaagtcattgcaatcgaagattttgaaacaggtaaaaag
actgacttacaaaatgatgaagtcagcgatataactttacctaaagcgaatgtaataaag
atatactttaatgaaggatttattgctttgcgtccttctggtacagagcctaaaattaaa
ctttatgtatcactttcttgtgaccattttgacgtagttgcacaaaaaatgaatgatgct
5       atatttaactcttaa


Sequence 2832
MIISYLHNINKLKNSNFLDLTQNKNGNMIYFKKEVILKMIDNWIDVLDESLVKDFYNNQT
SEEQQEGLDTTLSFGTAGIRGKFGLGEGRLNKFTVSKVALGFAHYLTSSIAHPVVVIHYD
10      TRHLSPDFAQIIANILASLDIKVYLADTYRTTPDLSFAVRYLQADAGIMITASHNPKDYN
GIKVYGEDGAQLSTDDSARLSTYIDKLGHPLHINLPSLTTEQQTLIHSVPSEVREDYFKN
VQDLVGTIPQSDLKVVFTSLHGTSVPVVPDILSSLNFNQFELVASQCEPDSDFSSVVSAN
PEDHKAFDQSIELANHIDADLLIGTDPDADRLGIVERDAEGNIHYYNGNQIGALLLNYRI
KQTEGLPNRIMFQSIVSGGLAKSLAQYHNVNFKEVLTGFKYIAAEIRHLSPEQNFIFGYE
15      ESYGFLARPFVRDKDAIQIVPLMIKYAAELKNKGRMLKDELEDITRNVGNFNDKLFSHTF
EGTQGKAKIENIMTQFRSETPSEMCGLKVIAIEDFETGKKTDLQNDEVSDITLPKANVIK
IYFNEGFIALRPSGTEPKIKLYVSLSCDHFDVVAQKMNDAIFNS*


Sequence 2833
20      Contig_0464_pos_2837_3322
is similar to (with p-value 4.0e-24)
 >sp:sp|P74561|HIS4_SYNY3 PHOSPHORIBOSYLFORMIMINO-5-AMINOIMI
DAZOLE CARBOXAMIDE RIBOTIDE ISOMERASE (EC 5.3.1.16). >gp:gp|
D90916|D90916_42 Synechocystis sp. PCC6803 complete genome,
25      26/27, 3270710-3418851. NID: g1653715.
gtgaccactaagcctatagaagtgggtggcggcattcgttcaaaacaaacaattgaaaat
tatattcattcaggaatagactattgtattgtaggtacaaaaggtatccaagatatagag
tggttaacacatatgacacatcaatttccaaataaactctacttatccgtagatgctttt
ggagagaaaataaagattaatggatggaaagaggatgctaaactcaatttatttgattat
30      gttgccaaaattgagcatttacctttgggtggtgtgatttataccgatatttcgaaagat
gggaaacttctggacctaattttgatttgacaggtcgtctcgcactttatacatcgttg
cctgtaattgcttcaggaggtattagacatcaagaggacttgtttcgattagaatcgtta
aatgttcatgctgctattgtaggaaaagcagcacatctggatgaattctgggagggatta
tcttga
35

Sequence 2834
VTTKPIEVGGGIRSKQTIENYIHSGIDYCIVGTKGIQDIEWLTHMTHQFPNKLYLSVDAF
GEKIKINGWKEDAKLNLFDYVAKIEHLPLGGVIYTDISKDGKLSGPNFDLTGRLALYTSL
·        PVIASGGIRHQEDLFRLESLNVHAAIVGKAAHLDEFWEGLS*
40

Sequence 2835
Contig_0464_pos_4161_4709
is similar to (with p-value 5.0e-40)
 >sp:sp|P44434|HIS2_HAEIN PHOSPHORIBOSYL-AMP CYCLOHYDROLASE
45      (EC 3.5.4.19) / PHOSPHORIBOSYL-ATP PYROPHOSPHOHYDROLASE (EC
3.6.1.31). >pir:pir|A64071|A64071 phosphoribosyl-AMP cyclohy
drolase (hisIE) homolog - Haemophilus influenzae (strain Rd
KW20) >gp:gp|U32730|U32730_6 Haemophilus influenzae Rd secti
on 45 of 163 of the complete genome. NID: g3212191.
50      atcaatgaagaagcttatcaaaaaactctgaaagaaaagaaagtaaccttcttctctaga
tctaacaacgtttatggactaaaggtgaaacttctggtcatttccaacacgttggagt
attcatctagattgtgatcaagatgcaatcttaatcaaagtgatgccacaaggtcctaca
tgtcacactggaagtctgagttgttttaatagtgaaattgaatcgcgctttaaaattcaa
gcattagcacaaacgattcatcaaagtgctaaaagcaatcaatctaactcttacactcaa
55      tatttattaaaggaaggtatcgagaaaatatccaagaaatttggtgaagaggcatttgaa
gttgtgataggtgcgataaaacataatcgtgaagaagttattaatgaaacagcagatgtc
atgtatcacctttttgtgttactacatagtttggatattccattttcagaagtagaacag
gtactagcgcatcgccatcaaaaaagaaataattttaaaggcgagcgcaaaaaggttcaa
gaatggtaa

Sequence 2836
MNEEAYCKTLKEKKVTFFSRSKQRLWTKGETSGHFQHVESIHLDCDQDAILIKVMIQGPT
CHTGSLSCFNSEIESRFKIQALAQTIHQSAKSNQSNSYTQYLLKEGIEKISKKFGEEAFE
VVIGAIKHNREEVINETADVMYHLFVLLHSLDIPFSEVEQVLAHRHQKRNNFKGERKKVQ
EW*

Sequence 2837
Contig_0465_pos_10122_9604
is similar to (with p-value 3.0e-49)
>gp:gp|Y13052|SSK3MECA1_3 S.sciuri mecA1 gene, strain K3(MM
2). NID: g2791901.
gtgagcatatatattatgaatatatgcgagtcattttatactataatcaatcgcgcacga
ttgaaaagaggtggagatgtgaaacaagaacaaatgaggttagcgaatcagctttgtttt
tcagcatataatgtaagtcgtttatttgctcaattttatgagaaaaagttaaaacagttt
ggtataacttattctcagtatttagtattactgacgttatgggaagagaatcctcaaaca
ttaaattcaattggtagacatttggatttatctagtaatactttaacccccttactaaaa
agacttgagcaatctggctgggttaaaagagaacgtcaacaatctgataaacgacagttg
ataattacgttaactgacaatgggcaacaacaacaagaagctgttttgaagcaatttca
agttgcttaccacaagaatttgatacgactgagtatgatgaaacgaaatatgtgtttgaa
gaactagagcaaacattaaaacatctcatagaaaaataa

Sequence 2838
VSIYIMNICESFYTIINRARLKRGGDVKQEQMRLANQLCFSAYNVSRLFAQFYEKKLKQF
GITYSQYLVLLTLWEENPQTLNSIGRHLDLSSNTLTPLLKRLEQSGWVKRERQQSDKRQL
IITLTDNGQQQQEAVFEAISSCLPQEFDTTEYDETKYVFEELEQTLKHLIEK*

Sequence 2839
Contig_0465_pos_9362_8244
>sp:sp|P47169|YJ9F_YEAST HYPOTHETICAL 161.2 KD PROTEIN IN N
MD5-HOM6 INTERGENIC REGION. >pir:pir|S57160|S57160 sulfite r
eductase homolog YJR137c - yeast (Saccharomyces cerevisiae)
>gp:gp|Z49637|SCYJR137C_1 S.cerevisiae chromosome X reading
frame ORF YJR137c. NID: g1015875.
gtgaattggaagaatttaatgcattacaaagaacaagtcattaatcctatgtctgaaacc
ctcacatcgatgtttgaacaacagggaattgatgtaatcatggggaaaggtaaacttgta
gatgctcatacaatagaggtaaataatacaactttacaatcagattatattgttatagca
actggacaacatagtcatcaattagatattgagggtaaagaatatacgcatgatagtcgg
gaatttttatcaatgcaatctttaccggatagtatcactttttattggagcaggtattatc
agtattgaattcgcttctatcatgatcaaatcaggtgtagaggttaatgtggttcatcat
acaaatcatgcacttgaagggtttaacgaatcacacgtcaataaattaattcaaaagtta
aaagatgaaggtgttaaattttactttagtgagaataccaagtcagttaaaccgaacgca
caacgttttatagtagaaactgagtctggaaagatgattgaaacagattatgtactggat
gcaaccggtagaaagcctaatgttcagcaaataggtttggaaaaagtgggtatactattt
agtgatagaggtattgaggttgacgattatttaagaacaaatgtgaaaaatatatacgca
agtggggacgttatcaataaaatgattcctaaacttactcctacagctacatttgagtct
aattatatcgctgcccatatccttggattgaatacagatgccattcagtatccaccaata
ccttcagtgctttattcattgcctcgtttatctcaaataggtgtcacagttagcgaggct
aagaaagatgatacgtatatgattaaagatataccattcggaagacaaatggtatttgag
tatcaaaacgaaacagaggctgaaatgtcaattgtattagatagtcacaaacgtttagta
ggagcagagatttatggtaatgacgctggtgatttggttaatctcctagtctttatcatt
aatcaaaaacttactgcacaagacttaaataaaaatattttttgcatttcctggagcttct
agtggcgttatagatttattgaaattggcaatgatgtag

Sequence 2840
VNWKNLMHYKEQVINPMSETLTSMFEQQGIDVIMGKGKLVDAHTIEVNNTTLQSDYIVIA
TGQHSHQLDIEGKEYTHDSREFLSMQSLPDSITFIGAGIISIEFASIMIKSGVEVNVVHH
TNHALEGFNESHVNKLIQKLKDEGVKFYFSENTKSVKPNAQRFIVETESGKMIETDYVLD
ATGRKPNVQQIGLEKVGILFSDRGIEVDDYLRTNVKNIYASGDVINKMIPKLTPTATFES

NYIAAHILGLNTDAIQYPPIPSVLYSLPRLSQIGVTVSEAKKDDTYMIKDIPFGRQMVFE
YQNETEAEMSIVLDSHKRLVGAEIYGNDAGDLVNLLVFIINQKLTAQDLNKNIFAFPGAS
SGVIDLLKLAMM*

Sequence 2841
Contig_0465_pos_8159_7755
>pir:pir|A34231|A34231 sulfite reductase (NADPH) (EC 1.8.1.
2) - Salmonella typhimurium >gp:gp|M23007|STYCYSJIHA_1 S.typ
himurium NADPH-sulfite reductase flavoprotein component (cys
J), NADPH-sulfite reductase hemoprotein component (cysI), an
d 3' phosphoadenosine 5'-phosphosulfate sulfotransferase (cy
sH) genes, complete cds. NID: g153928.
gtgctcattatcgtaaatactgcaacgaattgtactctgaatgatcaatttaataaatta
gaaacgcittataaaaaatatcataagtatggtcttgaaattttgagcttcccttgcaat
gattttaataatcaggagccagaattaatcaaagatatttatcgagtatataaatataag
tttggtattactttccccatccatgctaagattaatgttaatggagagcatgaacaccct
ttgtacacattattaaaatgtaaacaaccaggattatttggttcgcaaattaaatggaat
tttactaaatttgtagtagatcaacaggaaatattgttaaacgatttttaccttgtgat
aatcctaaccaaatggaaaaattaataagacaattattaaaataa

Sequence 2842
VLIIVNTATNCTLNDQFNKLEMLYKKYHKYGLEILSFPCNDFNNQEPELIKDIYRVYKYK
FGITFPIHAKINVNGEHEHPLYTLLKCKQPGLFGSQIKWNFTKFVVDQQGNIVKRFLPCD
NPNQMEKLIRQLLK*

Sequence 2843
Contig_0465_pos_6454_4568
is similar to (with p-value 1.0e-30)
>sp:sp|P52035|BSAA_BACSU GLUTATHIONE PEROXIDASE HOMOLOG BSA
A. >gp:gp|L77246|BACYACA_17 Bacillus subtilis (YAC10-9 clone
) DNA region between the serA and kdg loci. NID: g1256615. >
gp:gp|Z99115|BSUB0012_132 Bacillus subtilis complete genome
(section 12 of 21): from 2195541 to 2409220. NID: g2634478.
atgacaaatatcggaataacaaagtataagaggtggataagattgaacttaagtgttacc
aatagtccttttacggaagggcaagctaaacaaatcaatgaattgcttcacactttatca
tctaaccaacaagtatggttaagtggctatctgatggcaaatcaacaatcaaatacatct
acagattctgtagaacaacataactcagatgacaatacagaagcgatgctacatgaaaag
gaaccttcagttgagccggaagctagatcgataacaatttatatggatctgagtctggt
aatgcgcaaggacttgcagaaatatttgaacaacgtttatctgatattggaaatgacgtt
acgctcaaatcaatggatgactttaaacctaagaatttgaagaaagtagaagatttattt
attatcacatctacacatggcgaaggagatccaccagataacgctgttgaattacatgaa
tacatccacggacgtaaagcgccaaaattggatggggtgagattttcagtattagcatta
ggagatcaaacctatgaattcttctgtcaaactggtaaagattttgataaccgtttggct
gaattaggggcagaaagactttatcatcgtactgattgcgatgttgattatgaagaagac
gctgaaaagtggatggccaatgttattaatacgattgattcaacaccggctggtactgaa
agtgaacaagttgttagcgaatctataaaatcagcaaaagaaaagaaatatagtaagtct
aatccttacgatgctgaagtattgaccaatattaatttgaatggtagaggttcagacaaa
gagacgagacatattgagttattacttgataattttggtgaggaatatgaaccgggagat
tgtgtagttgtcttgcctcaaaatgatccagctatcgtagacttacttattagcacatta
ggttggagtccagaaacacaagtttttaattaacgaggatggagatactttaaatcttgaa
gaggcattaacatcgcattttgaaattactaaattaacaaaaccgttaatagaaaatgct
gcgatatttttgataatgaagagctttctgaaaaaattcaagataaagaatggattcaa
aactacgttgagggaagggatttgattgacttattaaatgacttcgcaacgacagaacta
caacctgaaaatttacatcaattattaagaaagttaccacctagagagtactcaacatct
agtagttataaagcaacaccagatgaagttcacattactgttggagcagttagatatcaa
gcacacggtcgggaacgttcaggtgtttgttcagtacaatttgcagagagaatacaagag
ggcgatacaattcctatctatttaaaacgaaatccgaattttaagtttccgcaagatgaa
tcaacacctgtgattatgataggtcctgggacaggtgttgcaccgtttagatcctatatg
caagaacgagaggaactaggttttgaaggaaatacatggttattctttggagatcaacac

```
          ttcactacagattttctgtatcaaacggaatggcaagaatggcttgaagatggaacttta
          tcaaaattagatgttgctttttctagagatactgataaaaaagtgtatgtgcaacataaa
          attgcagaaaatagtgaacaatttaatcgatggattgaaaatggcgctactatttatgta
          tgtggtgatgaaagtaaaatggcaaaggatgttcatcaagcgattaaaaatgtgttaatc
    5     aaagagcaaaacctatctgaaacagatgcagaagaatacttaaaacaaatgaaaagagat
          aaaagatatcaaagagacgtgtattaa
```

Sequence 2844

MTNIGITKYKRWIRLNLSVTNSPFTEGQAKQINELLHTLSSNQQVWLSGYLMANQQSNTS
TDSVEQHNSDDNTEAMLHEKEPSVEPEARSITILYGSESGNAQGLAEIFEQRLSDIGNDV
TLKSMDDFKPKNLKKVEDLFIITSTHGEGDPPDNAVELHEYIHGRKAPKLDGVRFSVLAL
GDQTYEFFCQTGKDFDNRLAELGAERLYHRTDCDVDYEEDAEKWMANVINTIDSTPAGTE
SEQVVSESIKSAKEKKYSKSNPYDAEVLTNINLNGRGSDKETRHIELLLDNFGEEYEPGD
CVVVLPQNDPAIVDLLISTLGWSPETQVLINEDGDTLNLEEALTSHFEITKLTKPLIENA
AIFFDNEELSEKIQDKEWIQNYVEGRDLIDLLNDFATTELQPENLHQLLRKLPPREYSIS
SSYKATPDEVHITVGAVRYQAHGRERSGVCSVQFAERIQEGDTIPIYLKRNPNFKFPQDE
STPVIMIGPGTGVAPFRSYMQEREELGFEGNTWLFFGDQHFTTDFLYQTEWQEWLEDGTL
SKLDVAFSRDTDKKVYVQHKIAENSEQFNRWIENGATIYVCGDESKMAKDVHQAIKNVLI
KEQNLSETDAEEYLKQMKRDKRYQRDVY*

Sequence 2845
Contig_0465_pos_4548_2830
 >gp:gp|Y13052|SSK3MECA1_4 S.sciuri mecA1 gene, strain K3(MM
2). NID: g2791901.

```
          atggttaatacaaataatcatatttcggaagaattagataaaaatcttgatgaaatggaa
          ttttttaaaagcaaatagtgactttttgcgtggaactattgaacaaagtttagctaatcca
          atcactggatccattacacaagatgatgcaaaactgctaaagtttcacggaagttatatg
          caagatgacagggatttaagagatgagcgtcgtaaacaaaaacttgagcctgcatatagt
          tttatgattcgagttcgtgtacctggggggaaagcgactcctgaacagtggattgctatg
          gatgatatctctaatcaatatgcaaatcatacgattaaattaacaacacgccaagcattt
          caatttcatggaattcttaaacgtaatttgaaacaatcaatgaaaaatattaatcatgca
          gtacttgattctattgctgcatgtggagatgttaatcgtaatacgatgtgcaatcctaat
          ccttatcaatctcaagtacataaggagattaatgattatgcaacgcgtataagtaatcac
          ttacttccaagaacaaatgcatatcatgaaatttggcttgatggtgaaaaggttttagat
   35     tcgagtgaggaaaaggaacctatttatgggaatacgtatttaccacgtaaattcaaaata
          ggtattgcagtaccaccatctaatgatattgacgtctattctcaagatattggtttaatc
          gctatcgttgaacaagatgagttaattggatttaatgtgactatcggtggcggtatgggt
          atgactcatggtaatactgaaacatatcctcaacttggacgtctcataggttttatacct
          aaggaaaaggttgtagatgtatgtgagaaatacttacaatacaacgtgattatggtaat
   40     cgtgaaaatcgaaaaaatgcacgttttaaatatacagtggaccgtctaggagaaacttgg
          gtgactgaagaattaaaccgacgattaggttgggaaattaaagcgccacgtgatttcgaa
          tttgaacataatggtgatcgattaggttggattgaaggtattaataattggaatttcact
          ttatttatacaaaatgggcgtgtgaaagatactgaagactatttgttaaaaacaacctta
          agagaaatcgcagaaatccatactggagatttcagattatcacctaatcagaacttagtt
   45     attgcaaatgtttctcctgagaaaaaggaagaaatacaagctattattgataaatataaa
          ttaacagatggcaaaaattatacaggacttagaagaaattctatggcttgtgttgctttc
          ccaacgtgtggtttagctatggcagaatctgaagatatcttccttcactaattacaaaa
          attgaagatttattagatgagtctggttttaaaagaggaagaaataacgattcgtatgaca
          ggttgtcccaatggatgtgcgagaccagcgctagcagaaatagcctttatcggtaaagca
   50     cctggtaaatataatatgtacttaggtggtagttttaaaggcgaacgtctaaataaaata
          tataaagagaatatcgacgaaaatgagatattagaaagtctacgtccattgttgttgcgt
          tatagtaaagagcgtcttgacggagaacactttgggggactttgtaattcgtgacggtgtg
          atagccaaagttcatgatggtcgcgattttcatagttaa
```

Sequence 2846
MVNTNNHISEELDKNLDEMEFLKANSDFLRGTIEQSLANPITGSITQDDAKLLKFHGSYM
QDDRDLRDERRKQKLEPAYSFMIRVRVPGGKATPEQWIAMDDISNQYANHTIKLTTRQAF
QFHGILKRNLKQSMKNINHAVLDSIAACGDVNRNTMCNPNPYQSQVHKEINDYATRISNH
LLPRTNAYHEIWLDGEKVLDSSEEKEPIYGNTYLPRKFKIGIAVPPSNDIDVYSQDIGLI

AIVEQDELIGFNVTIGGGMGMTHGNTETYPQLGRLIGFIPKEKVVDVCEKILTIQRDYGN
RENRKNARFKYTVDRLGETWVTEELNRRLGWEIKAPRDFEFEHNGDRLGWIEGINNWNFT
LFIQNGRVKDTEDYLLKTTLREIAEIHTGDFRLSPNQNLVIANVSPEKKEEIQAIIDKYK
LTDGKNYTGLRRNSMACVAFPTCGLAMAESERYLPSLITKIEDLLDESGLKEEEITIRMT
5   GCPNGCARPALAEIAFIGKAPGKYNMYLGGSFKGERLNKIYKENIDENEILESLRPLLLR
YSKERLDGEHFGDFVIRDGVIAKVHDGRDFHS*

Sequence 2847
Contig_0467_pos_7413_6943
10   is similar to (with p-value 2.0e-17)
     >sp:sp|P45637|YPRA_CORGL HYPOTHETICAL 33.0 KD PROTEIN IN PR
OB-PROA INTERGENIC REGION. >gp:gp|U31230|CGU31230_3 Coryneba
cterium glutamicum Obg protein homolog gene, partial cds, ga
mma glutamyl kinase (proB) gene, complete cds, and (unkdh) g
15   ene, complete cds. NID: g950194.
     atgaaagtaataggggttagtaagtcaggaaaaaatgttgaacaatttgatgaagtttat
     acaartgaagagttagatgatgttattgaaaaggcaaatattattgttaatgcattacct
     gaaacagaagaaacaatttacttattaaaaagaaaagatttcatacaaatggacaataat
     gccttatttataaatgtaggaagaggaacaattgttgatgaggaagtgctcatcaatgta
20   ctcaaagatcgattaatcagacatgcttatttagatgtttttgaaaaagaaccacttagt
     aaggacaatcctttatatgatttagataatgtgaccataactgctcatattacaggtaat
     gattctaataataatagagaagctacggacattttcaaaaagaatcttgagcattttctc
     aataattatgatgtaattgagaataaagtagacttagattatggttactaa

25   Sequence 2848
MKVIGVSKSGKNVEQFDEVYTIEELDDVIEKANIIVNALPETEETIYLLKRKDFIQMDNN
ALFINVGRGTIVDEEVLINVLKDRLIRHAYLDVFEKEPLSKDNPLYDLDNVTITAHITGN
DSNNNREATDIFKKNLEHFLNNYDVIENKVDLDYGY*

30   Sequence 2849
Contig_0468_pos_13714_9209
     >sp:sp|P39812|GLTB_BACSU GLUTAMATE SYNTHASE [NADPH] LARGE C
HAIN (EC 1.4.1.13) (NADPH-GOGAT). >gp:gp|Z99113|BSUB0010_138
Bacillus subtilis complete genome (section 10 of 21): from
35   1781201 to 2014980. NID: g2634090. >gp:gp|Z99114|BSUB0011_9
Bacillus subtilis complete genome (section 11 of 21): from 2
000171 to 2207900. NID: g2634230.
     gtgtgcatcatgtacaatgagaaattaaaaaaagggactatacgattatcgtgaagagcat
     gatgcgtgtggtattggattttatgccaatatggataataaaagatctcacgatattata
40   gaaaaatctttagaaatgttaagacggttagatcatcgtggtggagtaggtgccgatggt
     attactggtgatggtgcaggaattatgacggagataccataccaattgttcgaacaatta
     acagaattcaaagttcccggcgaaggatattatgccgtgggattattttttttctaaagag
     aaagttagagattcaattcacgaagagatgtttaatcaatattttgaaagtgaaggtttt
     aaagtcattggatatagagatgtgccagtagatactcgcgctattgctcaacatgttgca
45   gatactatgccttatattcaacaagtatttgttgacatcacaggtgtaaaagaagttgaa
     aaacgattgttttagcaagaaagcaaattgaaaatatagtgaaacacaatccatagat
     ttatattttacaagtctctctcatagaacgattgtttataaaggttggttacgttcggat
     caaattaaaggcttatatttagacctacaaaatgaggcatatcaatcaaaattaggactt
     gtacactcccgctttagtactaatacatttccaagttggaaacgtgcacatcccaatcgc
50   atgcttatgcacaatggtgaaattaataccattaagggtaacgtaaactggatgcgagca
     cgccaaaataaactagttgaaacattatttgaagatgagaaagataaggtgcattttatt
     gttgatgaagatggtagtgactcatcaatagttgataatgcgttagagttcttatcatta
     gcaatggagcctgaaaaagcagcgatgttattaattccagagccatggttatacaatgaa
     tctaacgataaaaaagttcgctctttctatgaattttatagttatttaatggagccatgg
55   gatggaccaactatgatttccttttgcaatggagataagataggtgcattgactgataga
     aatggattaagacctgggcgttatacaataactaaagacaatttattgtttttttcttcc
     gaagtaggtgtcattgatgttccagaagaaatgtagcatttaaaggacaacttaatcct
     ggaaagttattacttgtagacttttttgcaaaataaggttgtagaaaataatgagctaaaa
     actaatattgctaatgagttgccctacgaacaatggctaaagattataaaaataaaaat

697

```
gatttagataatatttattaccaatcttccgactgggatgatcaaacactcttccgctta
cagaaacaatttgcttacactaaagaagatatcaataaatatatgacagatttagtcatc
aataaaaaagatcccatcggagcgatgggatatgatgcacctattgcagttcttaatgat
aagcctgagtcactatttaattattttaaacaattatttgcacaagttaccaatccaccc
atcgatgcttatcgagaaaagattgtaactagtgaactttcatatttaggttcagaaggg
aacctcttatgtcctgatgaatcagttttagaaagaattcaattaaaaaaaccagtttta
aatgaagcgcaattatcatcaatcgatcattcgtattttaatgtaacgtatttatctaca
cttt?tacaggtgatttggaaagtagcttaaatgaactagggaaccgagcaataca?gct
gtacatcaaggtgcgaaaattttggtgttagacgatacgtctttaactcacgaaaatagt
tatgcaatgccaattttattagcgttaagtcacgtgcatcaattattaattcgagaagga
ttaagaatggagaccagcctcattgcgcagtccggtgaaacacgagaagttcatcacgtt
gcatgtttacttggttatggtgcaaacgctgtagttccatatttagcgcaacgaacgatt
gaacaattaacgcgtcaaggtcaactttcaggaactgtcgctgaaaatgttgctacgtat
accaatgtattgtcagaaggcgttattaaagtgatggctaaaatgggcatttctactgta
caaagttatcaaggagcacagatatttgaagcggtaggtttatcgaatagcgtcattgaa
aaatattttacaggtacacagtcaaaattatctggtataagtattgaacaaatagacaaa
gagaataaagcgagacaaagtgatgattctgattatcttgaatccggaagtgtattccaa
tggagacagcaaggtcagcatcatgcatttaatcctcgtacgatttttttattgcagcat
gcatgtagagaaaatgattacgagttatttaaaaaaattctccaaaactgtaaatttaaaa
cgtacggatcatattagacatttattagaattcaagacacgccaatctattgatattagt
cgtgttgaaccagcaagtgaaatcgtaaaaacgtttaatacaggagcaatgagttacggc
tctatctcagcagaggcacatgagacgttggctcaagctatgaatcaaattggaggtaaa
agtaatagtggagaaggtggagaagattcttcacgttacgaaattcaaaaggatggaagt
aataagataagtgcgattaagcaagttgcatcaggtcgttttggggtgacgagtgattac
ttgcaacatgcaaaagaaattcaaattaaagtcgcacaaggcgctaaaccaggggaaggt
ggacaactaccaggttcaaaagtatatccatggattgctgagactagaggttcgacacca
ggtataggattaattttcaccaccaccacaccatgatatttattcaattgaggacttagca
cagc?ca?tcatgatttaaaaaatgcaaatagaagagctgatattgcagttaagc?tcgta
tcaaaaactggcgttggaactatagcttcaggggtagctaaagctttcgccgataaaatt
gttataagtggttatgatggaggtacaggtgcatcgcctaaaacaagtattcaacatgca
ggtttgccatgggagataggccttgccgaaacacatcaaacacttaaattaaatgatttg
cgtagtcgcgtaaaattagaaacggatggtaagttactgacgggtaaagatgtagcttat
gcttgtgcgcttggtgcagaagaatttggtttcgcaacagcaccacttgttgttttggggg
tgtattatgatgagggtttgtcataacgatacgtgtccagtaggggttgcaacacaaaac
aaagatttaagagctttgtttagaggtaaggcacagcatgtagttaactttatgtatttt
atagctgaagaattacgtgaaattttggcttcacttggtttagaaacagtagaagagtta
gtaggaagaacagatcttcttcaacgttcgacgcaattgaaaccaaatagtaaagcagct
tcgcttcaaatagaacgtttaatagaacaatttgacggggttaatacgaaagagatatca
caaaaccatcatcttgatgaaggattcgatttgaattatctgtacccagacgcacgctat
agtattgaaaacgggcactcttttaccggaaattatgttgttaataatgaacagcgagat
gtaggtgtaattacaggtagtgcgatagctaaacaatatggagaagaaggattacctgaa
gatacgatacttgcttacactgaaggtcatgcaggtcaaagcttagctgcatatgcacca
cgcggattaacaatccatcataccggtgatgctaatgactacgtaggtaaaggattgtcc
ggtggaactgtcatcgtaaatgctccaaatagtcaacgtgaaaatgaaattatagcagga
aatataaacttttacggggcttctagaggtaaagcgtttatcaatggtaaagctggtgag
cgtttctgtatcagaaatagtggtgcagatgttgtagtagaaggtattggtgatcatgga
cttgaatatatgacaggggacatgtcattatcttaggagatgttggaaagaactt?ggc
caag?cat?gagcgggggcgtaagttatattttctcttctgacgtggagaaatttac?aaag
gttaatgcgcttgaaactttagaattcagtagcatacgttttgatgaggaaaaatctctt
atcaaagacatgcttgaagcacattttaagcatacacgtagtaacaaagcacgccaatta
cttgaccaatttgacaatattgaaaagttagcaattaaagttattccgaaagattacaaa
ttaatgatgcaaaaaattgatttgaaaaaacgtcaaatggaacgtgaagatgaagcaaca
ctggcagcgtttttatgatgacagagaaacaattgaacaagagctacagccagcagtcatt
tattaa
```

Sequence 2850
VCIMYNEKLKKGLYDYREEHDACGIGFYANMDNKRSHDIIEKSLEMLRRLDHRGGVGADG
ITGDGAGIMTEIPYQLFEQLTEFKVPGEGYYAVGLFFSKEKVRDSIHEEMFNQYFESEGF
KVIGYRDVPVDTRAIAQHVADTMPYIQQVFVDITGVKEVEKRLFLARKQIEKYSETQSID

698

LYFTSLSHRTIVYKGWLRSDQIKGLYLDLQNEAYQSKLGLVHSRFSTNTFPSWKRAHPNR
MLMHNGEINTIKGNVNWMRARQNKLVETLFEDEKDKVHFIVDEDGSDSSIVDNALEFLSL
AMEPEKAAMLLIPEPWLYNESNDKKVRSFYEFYSYLMEPWDGPTMISFCNGDKIGALTDR
NGLRPGRYTITKDNFIVFSSEVGVIDVPEENVAFKGQLNPGKLLLVDFLQNKVVENNELK

5   TNIANELPYEQWLKDYKNKNDLDNIYYQSSDWDDQTLFRLQKQFAYTKEDINKYMTDLVI
NKKDPIGAMGYDAPIAVLNDKPESLFNYFKQLFAQVTNPPIDAYREKIVTSELSYLGSEG
NLLCPDESVLERIQLKKPVLNEAQLSSIDHSYFNVTYLSTLYTGDLESSLNELGNRAIQA
VHEGAKILVLDDTSLTHENSYAMPILLALSHVHQLLIREGLRMETSLIAQSGETREVHHV
ACLLGYGANAVVPYLAQRTIEQLTRQGQLSGTVAENVATYTNVLSEGVIKVMAKMGISTV

10   QSYQGAQIFEAVGLSNSVIEKYFTGTQSKLSGISIEQIDKENKARQSDDSDYLESGSVFQ
WRQQGQHHAFNPRTIFLLQHACRENDYELFKKFSKTVNLKRTDHIRHLLEFKTRQSIDIS
RVEPASEIVKRFNTGAMSYGSISAEAHETLAQAMNQIGGKSNSGEGGEDSSRYEIQKDGS
NKISAIKQVASGRFGVTSDYLQHAKEIQIKVAQGAKPGEGGQLPGSKVYPWIAETRGSTP
GIGLISPPPHHDIYSIEDLAQLIHDLKNANRRADIAVKLVSKTGVGTIASGVAKAFADKI

15   VISGYDGGTGASPKTSIQHAGLPWEIGLAETHQTLKLNDLRSRVKLETDGKLLTGKDVAY
ACALGAEEFGFATAPLVVLGCIMMRVCHNDTCPVGVATQNKDLRALFRGKAQHVVNFMYF
IAEELREILASLGLETVEELVGRTDLLQRSTQLKPNSKAASLQIERLIEQFDGVNTKEIS
QNHHLDEGFDLNYLYPDARYSIENGHSFTGNYVVNNEQRDVGVITGSAIAKQYGEEGLPE
DTTLAYTEGHAGQSLAAYAPRGLTIHHTGDANDYVGKGLSGGTVIVNAPNSQRENEIIAG

20   NINFVGASRGKAFINGKAGERFCIRNSGADVVVEGIGDHGLEYMTGGHVIILGDVGKNFG
QGMSGGVSYIFSSDVEKFKKVNALETLEFSSIRFDEEKSLIKDMLEAHFKHTRSNKARQL
LDQFDNIEKLAIKVIPKDYKLMMQKIDLKKRQMEREDEATLAAFYDDRETIEQELQPAVI
Y*

25   Sequence 2851
Contig_0468_pos_9170_7728
is similar to (with p-value 5.0e-25)
 >sp:sp|P33019|YEIH_ECOLI HYPOTHETICAL 36.9 KD PROTEIN IN LY
SP-NFO INTERGENIC REGION. >gp:gp|U00007|ECOHU47_49 47 to 48

30   centisome region of E.coli K12 BHB2600. NID: g453983. >gp:gp
|AE000305|AE000305_3 Escherichia coli K-12 MG1655 section 19
5 of 400 of the complete genome. NID: g1788479.
atgaaatatgataaacagtcgctatcagaattgtctttggtagaccgtctttcgaatcat
gaagcgtttcaacaacgcttcactaaagaagatgcttcgattcagggtgcgcgctgtatg

35   gattgtggaacacctttttgtcaaactgggcaatcttatggaagagaaacaataggatgc
cctattggtaattatatacctgagtggaacgacttagtctatcatcaagatttttaaagct
gcttacgaaagattgagagagacgaataattttcctgaatttacaggaagagtttgtcct
gcaccatgtgagcaatcatgtgttatgaaattaatagagaatccgtggcgattaaaggt
att jaacgtacaattattgatgaagcatatgagaatgagtgggttcatcccgcatacct

40   gaagatcataaagaccaacgagttgctatcgtaggtagtggtccagcgggacttacagca
gctgaagaattaaactttaaaggctataaagttactgtttatgaaaaggcgcatgaacca
ggcggcttgctaatgtatggtataccaaatatgaaactagataaagacgtaatacgtcga
cgtgtatcacttatgaaagatgctggggtttttatttaaaacaggcgttgaaattggcgtc
gatgtgagccgtgaaacacttgaagaaattatgatgctattattttatgcacaggtgct

45   caaaatgcgagagatttaccattggaaggacgaatgggctctggtattcattttgcaatg
gactatcttactgaacaaacacagtatctaaatggtgagattgaaagtttgagcattact
gctaaagataagaatgtaattattataggtgctggtgatactggtgcagactgtgtagcg
acagcattacgtgaaaactgtaaatctattgttcaatttaataaatatacgaaacagcct
gaagagattactttgaaagtaatacttcctggccattagcaatgcctgttttcaaaatg

50   gattatgcgcataaggaatatgaagctaaatttggtcaagaaccaagagcctatggtgta
caaacaatgcgctatgatgttgacgagttaggaaatgttaaaggcttatatacacaaata
ttaaaagaaacgcctgatggcatggtgatggaagatggaccagaacgattttggccggct
gatttagtcttattatctatagggtttgttggtactgaaaccactgttccgcatgcgttt
gatatacacaccgagcgtaataaaattgtagctaatgatacaaattatcaaactaatcac

55   gctaaaatatttgctgcaggagatgcaagacgaggtcagagtttggttgtttgggcaata
aaagaaggtcgtgaagtagcacattctgttgatcaatacttaagtaaagaagttctagtg
taa

Sequence 2852

MKYDKQSLSELSLVDRLSNHEAFQQRFTKEDASIQGARCMDCGTPFCQTGQSYGRETIGC
PIGNYIPEWNDLVYHQDFKAAYERLRETNNFPEFTGRVCPAPCEQSCVMKINRESVAIKG
IERTIIDEAYENEWVHPAYPEDHKDQRVAIVGSGPAGLTAAEELNFKGYKVTVYEKAHEP
GGLLMYGIPNMKLDKDVIRRRVSLMKDAGVLFKTGVEIGVDVSRETLEENYDAIILCTGA
QNARDLPLEGRMGSGIHFAMDYLTEQTQYLNGEIESLSITAKDKNVIIIGAGDTGADCVA
TALRENCKSIVQFNKYTKQPEEITFESNTSWPLAMPVFKMDYAHKEYEAKFGQEPRAYGV
QTMRYDVDELGNVKGLYTQILKETPDGMVMEDGPERFWPADLVLLSIGFVGTETTVPHAF
DIHTERNKIVANDTNYQTNHAKIFAAGDARRGQSLVVWAIKEGREVAHSVDQYLSKEVLV
*

Sequence 2853
Contig_0468_pos_6592_5564
   >gp:gp|D85230|PEEGLTD_4 Plectonema boryanum URF141, ORF243,
   NADH-dependent glutamate synthase large subunit (gltB) and
   small subunit (gltD) and URF289 genes, partial and complete
   cds. NID: g1339947.
   gtgctaagatttcatatcgaggtgaaagtacatatgaaatcaataacgcaggcttcattt
   atgaaaggtattatgtttacatttacgattgcaataatcagttatatattagctaaattt
   cctattttacatacgattggggcgttagctattgccatcatttttgcgatgatataccgc
   caagtcataggttatcctgagcatattcgcccaggtattacgtttgcatcgaaacgttta
   ttaaaatttgcgattatcttatatgggttaaaattaaatatgggagatattctaggtaaa
   ggttggaaattactacttattgatattatcgtaattatcttttcaataagtttaactta
   cttttgaatcaaattattaaaggaaataaagatatctctatactacttggtattggtaca
   ggagtatgtggagctgcagctattgcagctacagcaccaatcttaaaatctaaagaaaaa
   gacattgcaataagtgtaggtattattgcactagttggaactatatttgcacttatttat
   acagctatcgaggctattttaacatacctactataacttatggtgcttggacaggtatc
   agtctacatgaaatcgctcaagtcgttttagcagcaggtattggcgggtcggaggcaatg
   acatttgctttacttggaaaattaggccgtgtgtttttacttattccattaagtattgtc
   ttgattttgtatatgcgttataagtcacactcaagtcaagtacaacaaaaaatcgatatt
   ccttactttctaattggatttattataatggcttgtatcaatacatttgttcctatt[?]cct
   tcatt[?]act[?]atgaatattataaatgttattacaacgttatgtatgttaatggcga[?]ggtt
   gctctaggattgaatatcgtttttaaaagaagttatttcaaaagcacttaaaccattcatt
   gtgatctgtataacttcaatttgtctgtctggtgtgactctcttagttacgtctataatg
   tttaaataa

Sequence 2854
VLRFHIEVKVHMKSITQASFMKGIMFTFTIAIISYILAKFPILHTIGALAIAIIFAMIYR
QVIGYPEHIRPGITFASKRLLKFAIILYGLKLNMGDILGKGWKLLLIDIIVIIFSISLTL
LLNQIIKGNKDISILLGIGTGVCGAAAIAATAPILKSKEKDIAISVGIIALVGTIFALIY
TAIEAIFNIPTITYGAWTGISLHEIAQVVLAAGIGGSEAMTFALLGKLGRVFLLIPLSIV
LILYMRYKSHSSQVQQKIDIPYFLIGFIIMACINTFVPIPSLLMNIINVITTLCMLMAMV
ALGLNIVLKEVISKALKPFIVICITSICLSGVTLLVTSIMFK*

Sequence 2855
Contig_0469_pos_5825_5232
is similar to (with p-value 2.0e-40)
   >gp:gp|Z67739|SPPARCETP_2 S.pneumoniae parC, parE and trans
   posase genes and unknown orf. NID: g1490398.
   atgttaatcttgagttatctgattggtgcattcccaagcgggttaattattggtaaatta
   ttt[?]tt[?]taaaaaagatataagacaatacggtagtggaaatactggagcaactaacag[?]ttt
   cgtgttc[?]tggaagaccagctggatttatagttacgttttttagatattttcaagg[?]attt
   attacagtcttttttccactatggttcccagttcatgcggatggtgttataagcaccttc
   tttacaaatggtttaatagtaggattgtttgcaatactcggtcacgtgtatccaatatat
   ctgaaatttaatggcggaaaagcagtagctaccagtgcaggagttgtattaggtgtcaat
   cctattttacttcttatcttggcaattatcttttttagtgtattaaaaatctttaaatat
   gtttctttatcaagtatcattgcagcaattagttgtgtgattggttcaatcatcattcat
   gattatattttacttgctgttagcggaattgtttcaatcatattaataattcgacacaaa
   tctaatatagttagaattttttaaaggagaagaacctaaaattaaatggatgtaa

Sequence 2856
MLTLSYLIGAFPSGLIIGKLFFKKDIRQYGSGNTGATNSFRVLGRPAGFIVTFLDIFKGF
ITVFFPLWFPVHADGVISTFFTNGLIVGLFAILGHVYPIYLKFNGGKAVATSAGVLLGVN
PILLLILAIIFFSVLKIFKYVSLSSIIAAISCVIGSIIIHDYILLAVSGIVSIILIIRHK
SNIVRIFKGEEPKIKWM*

Sequence 2857
Contig_0470_pos_7843_9180
is similar to (with p-value 3.0e-97)
  >gp:gp|Y13052|SSK3MECA1_1 S.sciuri mecA1 gene, strain K3(MM
2). NID: g2791901.
atgagtaaggttaatcatttgatagttgaagatgaacgttattttgcacattctggacgg
attaaatattatccactagttattgatcatggttatggagccacgttaatcgatgtagac
ggtaagtcttatattgatttattagcaagtgcaagttcgcaaaatgtgggtcacgctccg
aagccagtagtcgaagcaattaagaaccaaactgagaaattcattcattatacaccagca
tatatgtatcatgaaccgttagtgcgattatcaaaaaaattatgtgacattgctcctgga
aattatgagaagagagttactttcggattaagtggctctgatgctaatgatgggattata
aagtttgcacgagcgtatacaggacgtccatacatcataagtttactaatgcataccat
ggttcaacatttggttcattatcgatgtcttcaattagtttgaatatgcgtaagcattac
gggccattacttaatggattttaccatataccttttcccgataagtataggggatgttt
gagcaagctaagcctaacacggttgaggaatatttagctcctttaaaagaaatgtttgca
aaatacgtccctgcggaggaagttgcatgtattgtagttgaaacaattcaaggtgacggt
ggcttacttgaacccgtaccaggttattttgaagcattacaagagctttgccacgctcac
aatattcttattgcagttgatgatatccaacaaggattaggtcgtacaggaaagtggagt
tccgtagatcattattattttactccagatttaatgacatttggaaagtcattagctgga
ggtttaccaatgtctgcgattgtaggtcgtaaagaaatcatggaaagtcttgaagcacct
gctcatttatttacaactggtgcaaatcctgtaagttgcgaagcagccttagcaacgata
aagatgattgaagatgaagatttactaaacgcttcatggaaaaaggggagttacgttaga
aaaagaatagacccatggatagaacgttatcaatatgtaggtgatgttcgaggtattgga
ttatcgattggaatagacatagtatcaaataaaattgagaaaactagagattctgaagca
gcattaaagatatgtaattactgctttgaaaatggtgtgattatcatagcagttgcgggt
aatgtttttaagatttcaaccaccacttgtaattacctataagcaacttgataaagcatta
gatacaatagaacaggcgcttgaaaagttggaaagaggagaattaaatcaatatgacatt
agtggtcaaggttggtaa

Sequence 2858
MSKVNHLIVEDERYFAHSGRIKYYPLVIDHGYGATLIDVDGKSYIDLLASASSQNVGHAP
KPVVEAIKNQTEKFIHYTPAYMYHEPLVRLSKKLCDIAPGNYEKRVTFGLSGSDANDGII
KFARAYTGRPYIISFTNAYHGSTFGSLSMSSISLNMRKHYGPLLNGFYHIPFPDKYRGMF
EQAKPNTVEEYLAPLKEMFAKYVPAEEVACIVVETIQGDGGLLEPVPGYFEALQELCHAH
NILIAVDDIQQGLGRTGKWSSVDHYYFTPDLMTFGKSLAGGLPMSAIVGRKEIMESLEAP
AHLFTTGANPVSCEAALATIKMIEDEDLLNASWKKGSYVRKRIDPWIERYQYVGDVRGIG
LSIGIDIVSNKIEKTRDSEAALKICNYCFENGVIIIAVAGNVLRFQPPLVITYKQLDKAL
DTIEQALEKLERGELNQYDISGQGW*

Sequence 2859
Contig_0471_pos_8235_6703
  >gp:gp|Z46863|ACRBDOXN_10 Acinetobacter sp. cysD, cobQ, sod
M, lysS, rubA, rubB, estB, oxyR, ppk, mtgA, ORF2 and ORF3 ge
nes. NID: g2462044.
atggaagaaagaattggtttgatagacattggttccaacacgattcgacttgttatattt
ggctacaataaaaaaactgggctcaatgaaatactgaatataaaaacacctgcacgttta
agtcaatatctcactaagtccaatgaaatgaatgatgaaggtattcatgtttttaaaagag
actttaagcagttttagaaaagttgcggataaatttaacgttgatgcattatatcccatc
gcaacagctgctatccgtcaatctaaaaatcgtgaagctatcattaaagaaattaaacaa
gatattcatatcgaaattcaaattgtacctgaagaagatgaagcatttttacggttactat
gcgattacacatactactgatattgaaaatggaatttctgtcgatatcggtggcggttct
accgaagttacccttttcaaagacaaacaacttaaagaggctcatagctttccatttggc
gtggtatcacttaagcgtcagttttttggtgataaagaacataatgacaaaacagccatt

```
      aaaaatatggaacagtttttacgtgagcaatttagtcagttagactggctatccaaccaa
      catattgcgctcgtaggagtaggtggttctgcacgtaatgttgcacgcattcatcaatca
      gcacatgcataccctatcggtggcgttcataactatacgatgacttcgaaagatattaac
      aatgtttatgatctaatccgtaaaagttctcgcgatgaactaacaaatttagacggttta
 5    agtcgcgatcgcgtcgatattattctgccggctatctccgtctttaaaacacttttcaaa
      aaaattgacgccacacaattcaccttctcaagaaaaggaattcgtgaaggatttattatg
      aaccacatcagcaaacgatatcctgatgagtttaataaaagtaacgtacgtaaagatgca
      ttacgtcatttagcgaacgaatatcatattgaagaaacgagtgctaatcgtcgtgtgaaa
      ttagctcaatccttattgaatcaaattataagtgaacgatcacttaatatttcagaaatg
 10   gaaaaagaattatttattgaaggtgcctacatttattatctaggtagtttcattgattca
      gactcaagttcaccacatacgtattacttaatcgcaaattcaatgattaacggcttttca
      cataaagatcgtgtgaaattagctttgttagctagttttaaaaacaaatctttacttaaa
      ttttattgcaaagaaacacagtggtttagtaataaagaaatagatacaatacaagcttta
      ggagcgattattaaatttgcaaacaccttgaatatctcacatactagtttttgtagaggaa
 15   gttaaactaaaagcaaagaaagatgacaaatacgatttattagtttattacaaaggttca
      cctattgcagaagaataccaagcaaatcgtcagaaaaagcatattgagaaattttaaaa
      ggtaaggtttctattatatttacaaaatcttaa.
```

Sequence 2860

20   MEERIGLIDIGSNTIRLVIFGYNKKTGLNEILNIKTPARLSQYLTKSNEMNDEGIHVLKE
        TLSSFRKVADKFNVDALYPIATAAIRQSKNREAIIKEIKQDIHIEIQIVPEEDEAFYGYY
        AITHTTDIENGISVDIGGGSTEVTLFKDKQLKEAHSFPFGVVSLKRQFFGDKEHNDKTAI
        KNMEQFLREQFSQLDWLSNQHIALVGVGGSARNVARIHQSAHAYPIGGVHNYTMTSKDIN
        NVYDLIRKSSRDELTNLDGLSRDRVDIILPAISVFKTLFKKIDATQFTFSRKGIREGFIM
25   NHISKRYPDEFNKSNVRKDALRHLANEYHIEETSANRRVKLAQSLLNQIISERSLNISEM
        EKELFIEGAYIYYLGSFIDSDSSSPHTYYLIANSMINGFSHKDRVKLALLASFKNKSLLK
        FYCKETQWFSNKEIDTIQALGGIIKFANTLNISHTSFVEEVKLKAKKDDKYDLLVYYKGS
        PIAEEYQANRQKKHIEKILKGKVSIIFTKS*

30   Sequence 2861
        Contig_0471_pos_6659_4476
        is similar to (with p-value 1.0e-24)
          >gp:gp|AF083928|AF083928_2 Vibrio cholerae polyphosphate ki
        nase (ppk) and exopolyphosphatase (ppx) genes, complete cds.
35   NID: g3452464.

```
      atgaggtgtatgtatagtatgcaaactcgattgggagaaaaagatattaatttaccgcag
      tattacaacaataggggagttaagttggctagattttaactacagagtattacaagaatca
      tatgataaaaataatccgttgcttgaaaaacttaactttatttctatcttcagttcaaat
      ttagatgaattctttatggttcgagtggctgggttaaaagaccaagtcaaaatgggatat
 40   gacaaacctgaaaataaagcacagatgacgcctcaagaacaacttgatgctattaaaatt
      aaaaatactgactatgtgaacactcaatatcaacgttataacgaattaattaaagaatta
      gccaattacgatattgagatggtaaaacctgaagacttatcagatgcattgatagaaaaa
      ttagaaaagagttcaagttaagtgtcttgccgacactcactccgttaggtattgatgcg
      tatcatccatttccaaagttaaataataaaagtttaaatattttgttgatatcgatacg
 45   gaagatgccattaattcagctatcgttcaaattccttcattaattccacgctttttaact
      ttaaatgagggtacaaaacaatacgttgtcatggtagaagatgtaattacgtatttcatc
      aattatctatttacaggatacgaagtactaaatacttttactttccgaataacacgtaat
      gcagatttaaccattcatgaagatggcgctgaggacttgcttatagaaattgaacgtttc
      ttgaaagaacgtaagagtggttcggctgtacgtttagaattagattgtcgcacttctgaa
 50   aaagagaatgtagaatggttaatcgatcaattagaaattgaagataatgatatttattat
      ttagatggtccacttgatttaacattcttatttggattggttgatcatctatctcataag
      ctcaaatatttaacgtatgagaaatatacccctcaaccacctagatcattaggcaacaag
      aatatctatcaattatcattagaaagagatatattcttccaccatccgtatgaatcattt
      gaaccaatagttgactttattcgacaagcagcagatgacccaaatacaatcgctatcaaa
 55   caaaccttgtatcgagtgagtaaggattcgccgattattaacagcttaaaagaagccgct
      gaaaacggcaagcaagtaacggtgctcgtagaattaaaagcacgctttgatgaagaaaat
      aacgtacattgggcacgtatgttagaagatgctggctgtcacgttatttatggtatgaca
      catctaaaaacgcatagtaaaattgcgctagtcgttaaacgcatcaacaataaacttacg
      tcatttgttcatttaggcacaggtaactataacgataaaactgctaaattatacacagat
```

atgggtatcatcacgacaaataaagatatcgctgaggatgcaattaacttctttaattac
ttgagtggttactcaacaaaaccagaatacaataaattgattgtagcaccatacgatatt
cgagacgtcttcattgatcgtatcgataaagaaatacgtagtcatttacaacatggtaac
ggtaaaattatgatgaaaatgaactctttaaccgataaaacgattatcgaaaagctcttc
gaagcatcccaagcaggcgttaaaatacaactcatcattcgtggtatatgttgtcttaaa
ccaggcattccaggtattagcgaaaatatagaggttgttagtattgtgggtcgtttgctt
gaacattcacgtatttactacttccataataatggtgaggcgcatatttatttatcttca
gctgacgttatgacacgtaatatgattaaacgtgtcgaaatattattccctgttgaagat
aaatcaataggacaacgattagttaactatatgaatttacaattatctgacaaccaaaaa
ggtcgttaccaagatgcacaaggcgtttatcattatgtcgaaaacaattcatctcctta
aactctcaatcttacttgatgcaagaagcaattaagtatggagaagaactaaaaaaacaa
tcggtacaaccttctggacaacccgttcattctagacgtggtggtagttggatacgaaaa
ttaaaaagtacatttaaaagataa

Sequence 2862
MRCMYSMQTRLGEKDINLPQYYNNRELSWLDFNYRVLQESYDKNNPLLEKLNFISIFSSN
LDEFFMVRVAGLKDQVKMGYDKPENKAQMTPQEQLDAIKIKNTDYVNTQYQRYNELIKEL
ANYDIEMVKPEDLSDALIEKLEKEFKLSVLPTLTPLGIDAYHPFPKLNNKSLNIFVDIDT
EDAINSAIVQIPSLIPRFLTLNEGTKQYVVMVEDVITYFINYLFTGYEVLNTFTFRITRN
ADLTIHEDGAEDLLIEIERFLKERKSGSAVRLELDCRTSEKENVEWLIDQLEIEDNDIYY
LDGPLDLTFLFGLVDHLSHKLKYLTYEKYTPQPPRSLGNKNIYQLSLERDIFFHHPYESF
EPIVDFIRQAADDPNTIAIKQTLYRVSKDSPIINSLKEAAENGKQVTVLVELKARFDEEN
NVHWARMLEDAGCHVIYGMTHLKTHSKIALVVKRINNKLTSFVHLGTGNYNDKTAKLYTD
MGIITTNKDIAEDAINFFNYLSGYSTKPEYNKLIVAPYDIRDVFIDRIDKEIRSHLQHGN
GKIMMKMNSLTDKTIIEKLFEASQAGVKIQLIIRGICCLKPGIPGISENIEVVSIVGRLL
EHSRIYYFHNNGEAHIYLSSADVMTRNMIKRVEILFPVEDKSIGQRLVNYMNLQLSDNQK
GRYQDAQGVYHYVENNSSPLNSQSYLMQEAIKYGEELKKQSVQPSGQPVHSRRGGSWIRK
LKSTFKR*

Sequence 2863
Contig_0473_pos_5421_6053
is similar to (with p-value 8.0e-31)
 >gp:gp|D64024|D64024_2 Sulfolobus sp. DNA for 2-oxoacid:fer
redoxin oxidoreductase subunit alpha and beta, complete cds.
 NID: g1565182.
atggcaaacaaagatttaacagttatcgcttctggtggtgatggagacggctatgcaata
ggaatgggacatactattcatgctcttagacgtaatatgaatatgacgtatattgtcatg
gacaatcaaatatatggattaactaaaggacaaacatcaccttcctcagctaaaggattt
gtaactaaatcaacacctaaaggaaatatagaaaagaatgtagctccattggaattggca
ctgtcctctggtgcaacttttgtagcacaaggattctcaagtgatataaaggcattaact
aaaatgattgaagatgcgattcatcatgatggttttttcttttgttaatgtttttctcacct
tgtgttacttacaataaagtgaatacttatgactggtttaaagaacatttaacaagtatc
gatgatattgagggctatgacatcacagataaacaacttgctatgaaaactgtgctggat
catgagtcactggttaaaggtatcgtttatcaagatacaacaacaccttcttatgaatcg
caaattttcagaactagaacatgaggcgttagctaaaagagatattcatattacagaagaa
actttcaacgatttaactgcacaatttttataa

Sequence 2864
MANKDLTVIASGGDGDGYAIGMGHTIHALRRNMNMTYIVMDNQIYGLTKGQTSPSSAKGF
VTKSTPKGNIEKNVAPLELALSSGATFVAQGFSSDIKALTKMIEDAIHHDGFSFVNVFSP
CVTYNKVNTYDWFKEHLTSIDDIEGYDITDKQLAMKTVLDHESLVKGIVYQDTTTPSYES
QISELEHEALAKRDIHITEETFNDLTAQFL*

Sequence 2865
Contig_0476_pos_11219_12265
is similar to (with p-value 3.0e-44)
 >sp:sp|P36649|YACK_ECOLI PROBABLE 53.4 KD BLUE-COPPER PROTE
IN YACQ PRECURSOR. >gp:gp|AE000121|AE000121_8 Escherichia co
li K-12 MG1655 section 11 of 400 of the complete genome. NID

: g1786306.
atgtataataaagttttgcaattttaattataattttttccataataattattgcgtct
aatgatactttcgcagaaagtaagaatgatatgatgaatatgaaagaagataagaaaaat
acaatggatatgacaaatatgaaacatcatgacgaaagaaagaaattaaattcttcacaa
ggaaaaaatgaaataatatttcctaaagttgcagagtcaaaaaaagataacaatggttat
aaaaattatacattaaaagctcaggaaggaaagacagagttttacaaaaataattttтct
aatactctaggctacaatggaaatttacttggaccaactttaaaattaaaaaaaggagat
aaagttaaaattaagttaataaataacttagatgaaaatacaacatttcattggcatgga
ttagaagtaaatggaaaagtggatggagggccttctcaggttataaaaccaggaaaagaa
aaaactataaaatttgaggttaatcaagattctgctacgttatggtatcacccccacccc
tctccaaatacagctaaacaagtttataatggcttatcaggattattatatatagaagat
agtaaaaagaataattatcctagtgattatggaaaaaatgatttgcctataataatccaa
gataaaacatttgtatctaaaaaattaaattattcaaaaacgaaagacgaagatggcact
caaggtgatactgttcttgtgaacggaatagtaaaccccaaactgacaacaaaagaagag
aaaatacgtttgagacttttaaatggttctaatgctcgagatttaaatcttaagctaagt
aataatcaaagttttgagtatattgcttcagatggcggtcaattaaaaaaacgctaaaaaa
ttaaaagaaattaatttagctccttcagaaagaaaagaaatagtaatagatttatctaaa
atgaaaggcgagaaaatcagtctggttgataatgataaaactgtaattttaccgattagt
aaacaactggtaagtgtaagatattag

Sequence 2866
MYNKVFAILIIIFSIIIIASNDTFAESKNDMMNMKEDKKNTMDMTNMKHHDERKKLNSSQ
GKNEIIFPKVAESKKDNNGYKNYTLKAQEGKTEFYKNNFSNTLGYNGNLLGPTLKLKKGD
KVKIKLINNLDENTTFHWHGLEVNGKVDGGPSQVIKPGKEKTIKFEVNQDSATLWYHPHP
SPNTAKQVYNGLSGLLYIEDSKKNYPSDYGKNDLPIIIQDKTFVSKKLNYSKTKDEDGT
QGDTVLVNGIVNPKLTTKEEKIRLRLLNGSNARDLNLKLSNNQSFEYIASDGGQLKNAKK
LKEINLAPSERKEIVIDLSKMKGEKISLVDNDKTVILPISKQLVSVRY*

Sequence 2867
Contig_0476_pos_6429_5413
is similar to (with p-value 6.0e-55)
 >gp:gp|X97452|ECPAA_12 E.coli paa cluster for phenylacetic
acid degradation. NID: g2764820.
atgtctcttttggatgacgtcatttttaggtaatacggtaggtaatgggggggaatttagct
agaaaatcattacttgaagcgggattagatttttaaaatacctggtataacaattgatcgt
caatgtggctcaggtcttgaagccgttatacaagcctgtaggatggtacaaagtgqtgct
ggaacaatatatattgcaggtggtgttgagagtaccagtagagcaccttggaaaatcaaa
cgtccgcagtcagtttatgaatctgagtttccacaatttttgaacgggcgccttttgca
agagaaggagaagacccttcaatgattgaagcagccgaaatgtagcgaagaaatatcat
atcagtagaaatgaacaagatgactttgcgtatcgcagtcatcagttggcatcaaaaaat
atgaataacggtaatatttcccaagaaattttaccgttcaaagtgaaaggtgaatatttt
aatcaagatgaaagtattaaacctcaacttactctcagaacacttggcagacttaaacca
cttttaaatgaaggaacagtcacagtaggaaatagttgtaagaaaatgatggtgcagta
ttactgattgttatggaagaaatcgggcacgtcaattaggattcacagaagggattaag
tttgtgaatagtgcaactgtaggtgttcaaccacagtatttaggagtaggtccagtgcca
gcagtaaatcaattattagctcaagaacgattaactataaatgatataaatgcagtagaa
ttaaatgaagcatttagctctcaagttattgcgagccaacaacagcttaacattcctttg
aataagttgaattgttggggaggagcaattgctacagggcatccatatggtgcaagtgga
gcagcgttagtcacacgtttatttatatgaaacatcaatttagaactatagcaactatg
ggaataggtggagggataggaaatgcagctttatttgaaagatggtatggaaattag

Sequence 2868
MSLLDDVILGNTVGNGGNLARKSLLEAGLDFKIPGITIDRQCGSGLEAVIQACRMVQSGA
GTIYIAGGVESTSRAPWKIKRPQSVYESEFPQFFERAPFAREGEDPSMIEAAENVAKKYH
ISRNEQDDFAYRSHQLASKNMNNGNISQEILPFKVKGEYFNQDESIKPQLTLRTLGRLKP
LLIEGTVTVGNSCKKNDGAVLLIVMEENRARQLGFTEGIKFVNSATVGVQPQYLGVGPVP
AVNQLLAQERLTINDINAVELNEAFSSQVIASQQQLNIPLNKLNCWGGAIATGHPYGASG
AALVTRLFYMKHQFRTIATMGIGGGIGNAALFERWYGN*

704

Sequence 2869
Contig_0477_pos_254_727
is similar to (with p-value 8.0e-35)
>gp:gp|U96107|SCU96107_3 Staphylococcus carnosus N5,N10-met
hylenetetrahydromethanopterin reductase homolog, SceB precur
sor (sceB) and putative transmembrane protein genes, complet
e cds, and putative Na+/H+ antiporter NhaC (nhaC) gene, part
ial cds. NID: g2735503.
atgaaaaaatcaaaacaatctcgacattggtagctggacttggtatagcatttctaggt
cacacaacacatgcagatgcggctgaaaataacaatcaacaacaaagtacatataactat
agtacaactgaagtatcattttctaattcaggaaatttatatacttctggccaatgtact
tggtatgtttatgataaaactggtggaaaaatcggatcaacatgggggaatgcaaatagc
tgggcaactgcagctcaagcagcaggattcactgtaaataatacacctgaagaaggtgca
attatgcaatcatctgaaggtgctttcggacatgttgctttcgttgaaagtgtcaataat
gatggttctattactgtatcagaaatgaactatgatggtggtccattcgctataagcaca
cgaacaatctctgccagtgaagcaagttcatataattacatccacctgaattaa

Sequence 2870
MKKIKTISTLVAGLGIAFLGHTTHADAAENNNQQQSTYNYSTTEVSFSNSGNLYTSGQCT
WYVYDKTGGKIGSTWGNANSWATAAQAAGFTVNNTPEEGAIMQSSEGAFGHVAFVESVNN
DGSITVSEMNYDGGPFAISTRTISASEASSYNYIHLN*

Sequence 2871
Contig_0478_pos_6998_6654
is similar to (with p-value 1.0e-60)
>pir:pir|I67760|I67760 transposase (insertion sequence IS10
) - Escherichia coli >gp:gp|S67119|S67119_2 BST=somatotropin
...BST/beta-Gal fusion protein [Escherichia coli, LBB84 pla
smid pXT107, IS10L/R-1, PlasmidTransposonInsertionMutant, 3
genes, 1679 nt]. NID: g455674.
atgcagattgaagaaaccttccgagacttgaaaagtcctgcctacggactaggcctacgc
catagccgaacgagcagctcagagcgttttgatatcatgctgctaatcgccctgatgctt
caactaacatgttggcttgcgggcgttcatgctcagaaacaaggttgggacaagcacttc
caggctaacacagtcagaaatcgaaacgtactctcaacagttcgcttaggcatggaagtt
ttgcggcattctggctacacaataacaagggaagacttactcgtggctgcaaccctacta
gctcaaaatttattcacacatggttacgctttggggaaattatga

Sequence 2872
MQIEETFRDLKSPAYGLGLRHSRTSSSERFDIMLLIALMLQLTCWLAGVHAQKQGWDKHF
QANTVRNRNVLSTVRLGMEVLRHSGYTITREDLLVAATLLAQNLFTHGYALGKL*

Sequence 2873
Contig_0480_pos_5372_6550
is similar to (with p-value 2.0e-24)
>sp:sp|P23524|YHAD_ECOLI HYPOTHETICAL 42.1 KD PROTEIN IN RN
PB-SOHA INTERGENIC REGION (ORF 3) (F408). >pir:pir|JQ0614|JQ
0614 hypothetical 42K protein - Escherichia coli >gp:gp|D902
12|ECORNPBW_3 E.coli rnpB gene and ORFs. NID: g216630. >gp:g
p|U18997|ECOUW67_54 Escherichia coli K-12 chromosomal region
from 67.4 to 76.0 minutes. NID: g606010. >gp:gp|AE000394|AE
000394_2 Escherichia coli K-12 MG1655 section 284 of 400 of
the complete genome. NID: g2367197.
atgtttaaaataatttttggaaaagagaaaaataaggtggttaagacaatgaaagtttta
gtagccatggatgaatttaatggaattatttctagttaccaagctaatagatatgttgaa
gaagcggtagcaagtcaaattgaagatgcagatatcgttcaagttccactatttaacggt
cgtcacgaattattagattcagtctttctttggcaatcaggaaataaatatcgtgtgagt
gcgcatgatgctgacatgaaagaaaccgaagcaatatatggacaaacggatagtggtatg
actattatcgaaggtcacttatttttaaatggcaaaaaacctattcaacatcgatcaagt
tacggtttgggagaggttataaaagcagcattggacaatcatacagaacatcttgttatt

```
      tctttaggtggaataggaagttttgatggcggtgcaggcatgttgcaagcattgggtgca
      acattttatgatgatgaagcacaaattgtcgatatgaggaaaggtgcatatttaataaaa
      tatattagacgtattgatttatcaggtgttcatccacaattaacaaaggtaaacattcaa
      ttaatgtcagatttctcaagtcgattgtatgggaaaaaaagtgaaatcatgcaaacatac
 5    gaatcattagatttgtctcaaaatgaagcagccgagatagataatttaatttggtatttt
      agtgaattatttaagaatgaattgaaaatagcaatgggaccaatcgagcgcggtggtgct
      ggaggtggtatagcagctgtattaaatagtctctatcaagctgagattttaacaagccat
      gaattagtgaatcaaatcacacatttagaaaacttaattcaacaggcagatcttattatt
      ttcggagaaggttttgaaagaagaagatcaaattctagagactacaacaatacgtatagca
 10   gaacttacccaacaatacagcaagccagctattgcaatttgtgctacaaatgataaattt
      gatttgtttgaatcattgaatgttacagcaatgtttaatacatttattgatatgcctgat
      tcatatacagattttaagatgggtattcaaatcagacattacacagtacaagcactaaaa
      ttattgaaaacgcaaataaacttaccgctttcatcctaa
```

15 Sequence 2874
    MFKIIFGKEKNKVVKTMKVLVAMDEFNGIISSYQANRYVEEAVASQIEDADIVQVPLFNG
    RHELLDSVFLWQSGNKYRVSAHDADMKETEAIYGQTDSGMTIIEGHLFLNGKKPIQHRSS
    YGLGEVIKAALDNHTEHLVISLGGIGSFDGGAGMLQALGATFYDDEAQIVDMRKGAYLIK
    YIRRIDLSGVHPQLTKVNIQLMSDFSSRLYGKKSEIMQTYESLDLSQNEAAEIDNLIWYF
20 SELFKNELKIAMGPIERGGAGGGIAAVLNSLYQAEILTSHELVNQITHLENLIQQ3DLII
    FGEGLKEEDQILETTTIRIAELTQQYSKPAIAICATNDKFDLFESLNVTAMFNTFIDMPD
    SYTDFKMGIQIRHYTVQALKLLKTQINLPLSS*

    Sequence 2875
25 Contig_0483_pos_5289_6392
    >pir:pir|G64047|G64047 cystathionine gamma-synthase (metB)
    homolog - Haemophilus influenzae (strain Rd KW20) >gp:gp|U32
    694|U32694_2 Haemophilus influenzae Rd section 9 of 163 of t
    he complete genome. NID: g1573035.
30 atgaaggatacagatttagctcaaattgctttaacacaagatcacactggcgcaattgcc
    aatccaatatatttatctactgcatatcagcatcctcacctaggtgaatcaacaggctat
    gattatacacgaactaaaaatccaacaagaacggcctttgaagaagcttttgcacaactt
    gaaaaaggcattgcttcatttgctacttccagtggtatggcggcgattcagttaatatgt
    aatatattcaaaccaggtgatgaaattctcgttgcatttgacctatatggtggaacattt
35 cggttattcgattttttacgaaaaacaatatggtttgaagtttaaatatgtagactttta
    aattatgaagaagtggaaaaaaacatcactccacaaactagagcattatttattgaacca
    atctcaaatccacaaatgattgaaattgatgtagaaccatattatatccttagcaaaaaa
    catcaactattaacaattatcgacaacacttttttaacaccttatctttcgacaccactc
    gaagaagptgcagatatcgttctgcattcagcaacaaaatatattggcggacataacgat
40 gtgttagctggagttgtaactgttaaggatgctcaattagctgaacaattgaatcaattc
    cataatatgattggagcaactctatcacctcttgatagttatcttttacaaagaggtcta
    aagacattacatcttcgcatagagcgttcccaagaaaatgctcaaaaacttgcacaacga
    tgtcgccagtcagattcaattgatgaagtttatatagtggacgaacaggcatgcttagt
    ttaagactgaaccaagcatatagcgtcgctaaatttttagaaaatttagaaatttgtata
45 tttgcagaaagcttaggtggtacagaaacatttatcaccttcccttatacacaaacacac
    gttgatatgccagatgaggaaaaagacaaacgtggaattgatgaatatctcatcagattg
    tccgtaggtatagaagactataacgatatagaagctgacataattcaagcattagagaat
    tctaaagtaggagtgatttcatga

50 Sequence 2876
    MKDTDLAQIALTQDHTGAIANPIYLSTAYQHPHLGESTGYDYTRTKNPTRTAFEEAFAQL
    EKGIASFATSSGMAAIQLICNIFKPGDEILVAFDLYGGTFRLFDFYEKQYGLKFKYVDFL
    NYEEVEKNITPQTRALFIEPISNPQMIEIDVEPYYILSKKHQLLTIIDNTFLTPYLSTPL
    EEGADIVLHSATKYIGGHNDVLAGVVTVKDAQLAEQLNQFHNMIGATLSPLDSYLLQRGL
55 KTLHLRIERSQENAQKLAQRCRQSDSIDEVLYSGRTGMLSLRLNQAYSVAKFLENLEICI
    FAESLGGTETFITFPYTQTHVDMPDEEKDKRGIDEYLIRLSVGIEDYNDIEADIIQALEN
    SKVGVIS*

    Sequence 2877
```

Contig_0484_pos_4823_4407
>sp:sp|O69282|MQO_CORGL MALATE:QUINONE OXIDOREDUCTASE (EC 1
.1.99.16) (MALATE DEHYDROGENASE (ACCEPTOR)) (MQO).
atgataggtactatgattgaaacgcctagagcatgcttgattgcgaatgaccttgcgaaa
cattgtgatttcttcagttttggtactaatgatttaacgcaattgacatttggtttctct
agagatgatgcaggaaaattcataaatgtgtatactgaaaataacattttacagcttgac
ccattccaaactttagatagagaaggtgtaggacgactaattcaattagctgttgaacaa
gctaaaaatacaaatccagagataaaaattggtgtatgtggtgagcttggtggcgatgca
aaatcaattcgtaaatttaaccaatgggaaattgattatgtttcatgttcaccatttaga
gttccgggtgcaattttagctacagctcagagtcaagcggaggaaagcgagcgataa

Sequence 2878
MIGTMIETPRACLIANDLAKHCDFFSFGTNDLTQLTFGFSRDDAGKFINVYTENNILQLD
PFQTLDREGVGRLIQLAVEQAKNTNPEIKIGVCGELGGDAKSIRKFNQWEIDYVSCSPFR
VPGAILATAQSQAEESER*

Sequence 2879
Contig_0484_pos_3408_1879
is similar to (with p-value 6.0e-42)
>pir:pir|S53297|S53297 pyruvate,orthophosphate dikinase (EC
2.7.9.1) - Flaveria pringlei >gp:gp|X75516|FPPDK_1 F.pringl
ei mRNA for pyruvate, orthophosphate dikinase. NID: g577775.

atgataactactacacaagaggtgaacattatggctatgtctgacaaaaaagacgtcgtg
ttaatcggtgctggtgtactaagtactacatttggttctatgttgaaaacgattgcacct
gattgggacattcatttatatgaacgtctagatcgtcctggtattgaaagttcaaatgaa
cgtaacaatgcaggaacaggacatgcagctttatgtgaattgaactatactgtacaacaa
cctgatggttcaattgatattgaaaaagctaaagaaattaatgaacaatttgaaatttct
aaacaattctggggtcatttagttaaatcaggagaaattcaaaatcctaaagaatttatt
aatccattacctcatattagttttgttcgtggtaaaaataacgttaaattcttaaaagat
cgttatgaagcgatgaagcaattccctatgttcgataatatcgaatatactgaagatatt
gaagaaatgagaaaatggattccattaatgatgaaaggccgtgaagataagggctacatg
gcagcgagtaaaatagacgaaggaactgacgtaaactacggtgaattaactcgtaaaatg
gctcaaatcttaaaaactcaccaaacgttgaagtgcaatacaaacatgaagttgttgat
tttgaacgtttgtctaatggtaaatggtcagttaaaattaaaaatctaaataatggacaa
gtattcgaacatcaaactgattatgtgtatttatcggtgctggtggtggcgcaattccacta
ttacaaaaaactggcattccagaaagtaaacatttaggtggattcccaatcagtggtcaa
tttattgcttgtacaaatccgcaagttattgaacaacacgatgccaaagtttatggtaaa
gaaccacctggtacaccaccaatgacggtacctcacttagatacacgttacattgatggt
gaaagaacattattatttggaccatttgcaaacgtgggacctaaattcctaaaacatggt
tctaacttggatttattcaaatcaattaaaccatataacattacgactttacttgcttct
gcagttaaaaacttaccattaattaagtattcatttgaccaagtcatcatgacaaaagaa
ggttgtatgaaccatttacgtacgttctatcctgaagcacgtgatgaagattggcaagtt
tatacagctggtaaacgtgtacaagttattaaagatactgaagagaatggtaaaggattt
atccaatttggtactgaagtggttaattctgaagaccactcagttattgcattactaggc
gaatcaccaggagcatcaacttcagtatcagtagcactagaagtttttagagaaaaacttc
cctgaatacgcgaaagactgggaacctaaaatcaagaaaatgattccttcatatggtgaa
tcacttattgatgatgttcaattaatgagaaaaatacgtaaacaaacatctaaagatctt
gaattaggattctataataaagcaaaataa

Sequence 2880
MITTTQEVNIMAMSDKKDVVLIGAGVLSTTFGSMLKTIAPDWDIHLYERLDRPGIESSNE
RNNAGTGHAALCELNYTVQQPDGSIDIEKAKEINEQFEISKQFWGHLVKSGEIQNPKEFI
NPLPHISFVRGKNNVKFLKDRYEAMKQFPMFDNIEYTEDIEEMRKWIPLMMKGREDKGYM
AASKIDEGTDVNYGELTRKMAQNLKNSPNVEVQYKHEVVDFERLSNGKWSVKIKNLNNGQ
VFEHQTDYVFIGAGGGAIPLLQKTGIPESKHLGGFPISGQFIACTNPQVIEQHDAKVYGK
EPPGTPPMTVPHLDTRYIDGERTLLFGPFANVGPKFLKHGSNLDLFKSIKPYNITTLLAS
AVKNLPLIKYSFDQVIMTKEGCMNHLRTFYPEARDEDWQVYTAGKRVQVIKDTEENGKGF
IQFGTEVVNSEDHSVIALLGESPGASTSVSVALEVLEKNFPEYAKDWEPKIKKMIPSYGE

SLIDDVQLMRKIRKQTSKDLELGFYNKAK*

Sequence 2881
Contig_0486_pos_1073_1840
is similar to (with p-value 9.0e-58)
  >gp:gp|U38892|SSU38892_2 Synechocystis sp. ruvB gene, compl
ete cds, and secA gene, partial cds. NID: g1256587.
atgtttaaaataggaaatttagaattacaatctcgtttacttttaggtactggaaaattt
gaaatgaagaggttcagtcaaaagcaattgaggcatctgaaacaaatgtacttacattt
gcagtaagacgtatgaatttatatgatcgtaacctacctaacccacttgcaaacgttaat
ttaaaagattttatcacttttccaaatactgcaggtgccaaaacagctcaagaagctatc
agaattgctgaaattgctaatcacgcaggtgtatgtgacatgattaaagtcgaagtcatt
ggtgatgacgaaacattattacctgatccattcgaaacatacgaggcatgcaaagtattg
ttagaaaaggttacactgtttgtccttacatctctaacgatttagttttagctcaacgt
ttagaagaattgggtgtacacgcagttatgccacttgcatcccctattggtacaggaaga
ggtattaataacccattaaatttaagttatattatcgaaatgctagtgtacctgtaatc
gtagatgctggtattggttcccctaaagatgcgtgtcatgccatggagcttggcgcagat
ggtattttactcaacacagccatttcagcggcaaaagatcctgtgaaaatggctgaagca
atgaaattaggtataaatgctggcagactttcatatgaagctggacgcattcctgttaag
tatactgcacaagcatctagtccatcagaaggtttagggttcttgtaa

Sequence 2882
MFKIGNLELQSRLLLGTGKFENEEVQSKAIEASETNVLTFAVRRMNLYDRNLPNPLANVN
LKDFITFPNTAGAKTAQEAIRIAEIANHAGVCDMIKVEVIGDDETLLPDPFETYEACKVL
LEKGYTVCPYISNDLVLAQRLEELGVHAVMPLASPIGTGRGINNPLNLSYIIENASVPVI
VDAGIGSPKDACHAMELGADGILLNTAISAAKDPVKMAEAMKLGINAGRLSYEAGRIPVK
YTAQASSPSEGLGFL*

Sequence 2883
Contig_0487_pos_1335_2006 ,
is similar to (with p-value 5.0e-34)
  >sp:sp|P39762|AMPS_BACSU AMINOPEPTIDASE AMPS (EC 3.4.11.-).
  >gp:gp|AF012285|AF012285_21 Bacillus subtilis mobA-nprE gen
e region. NID: g3282109. >gp:gp|Z99111|BSUB0008_117 Bacillus
subtilis complete genome (section 8 of 21): from 1394791 to
1603020. NID: g2633699.
atgacgaattatcataataagttaaaacaatatgcagaattattagtaagagtgggaatg
aatgtacaaccacagcaacctgttttttatacgttcatctgttgaagcgttagaattaact
catttaatcgtcgaggaagcatataaagcaggggcagaagatgttcgagtgagctacaca
gacccgaaattaaaaagattaaaatttgaaaacgaatcagttgaacactttgaaaaacaa
gaactcaaacaatatgatattgaagagcgtctggattatgttaatcgtggcgcagcgaac
ttggcgctcattgctgaagatccagagctattaaatggaatagatgcgcaaaagttaaaa
gcgtatcaaactgtatactcaaaaggatttaaaccatatatggaagcaagtcaaaaaaac
caatttccatgggtagtggctgcgttccctactagggattgggcacgtcgtgtctatcca
gagttggatgttgaatcagcatatattaaattcattgatgaagtatttgatattgttcgt
gtagatggacaaaatccaattgaaaattgggaaaaacacattaaagattaagtgttcat
gctaaacgattacaagagaaaaactatcaagctttacattacatatcagaaaattcatac
atttggttttga

Sequence 2884
MTNYHNKLKQYAELLVRVGMNVQPQQPVFIRSSVEALELTHLIVEEAYKAGAEDVRVSYT
DPKLKRLKFENESVEHFEKQELKQYDIEERLDYVNRGAANLALIAEDPELLNGIDAQKLK
AYQTVYSKGFKPYMEASQKNQFPWVVAAFPTRDWARRVYPELDVESAYIKFIDEVFDIVR
VDGQNPIENWEKHIKDLSVHAKRLQEKNYQALHYISENSYIWF*

Sequence 2885
Contig_0487_pos_3511_3086
is similar to (with p-value 6.0e-23)
  >sp:sp|P14597|DUT_ORFN2 DEOXYURIDINE 5'-TRIPHOSPHATE NUCLEO

TIDOHYDROLASE (EC 3.6.1.23) (DUTPASE) (DUTP PYROPHOSPHATASE)
. >gp:gp|M30023|ORFPRTPS_1 Orf virus homologue of retroviral
pseudoprotease gene, complete cds. NID: g332561.

atgacaaatacattagaaattaaattattatcagaaaacgcgactatgccgaagagagca
aattctacagatagtggattggacttatacgtatcagaaacgattaacattcctgcacac
gcaactaaagtagttaaaacagatatagcgattaatctgccttatgggtatgaggcgcaa
gtaagacctagatctggtaaatcacttaaaactaaattgcgtgtagcactaggaacaata
gaccaaacataccacaaagaaataggtatcatcacagataacataggtaatgaagatatc
acagtagaaaaaggagaaagattagcgcaattagttgtagcgccagttgtatatcctaca
cccaaacaggttgattggtttgaaaatgaaagcgacagaggtgcatatggaagcacagga
gaataa

Sequence 2886
MTI!TL3I!;LLSENATMPKRANSTDSGLDLYVSETINIPAHATKVVKTDIAINLPYGYEAQ
VRPRSGhSLKTKLRVALGTIDQTYHKEIGIITDNIGNEDITVEKGERLAQLVVAPVVYPT
PKQVDWFENESDRGAYGSTGE*

Sequence 2887
Contig_0487_pos_988_515
is similar to (with p-value 2.0e-17)
>sp:sp|P41893|PPAL_SCHPO LOW MOLECULAR WEIGHT PHOSPHOTYROSI
NE PROTEIN PHOSPHATASE (EC 3.1.3.48) (LOW MOLECULAR WEIGHT C
YTOSOLIC ACID PHOSPHATASE) (EC 3.1.3.2) (PTPASE) (SMALL TYRO
SINE PHOSPHATASE). >pir:pir|A55446|A55446 protein-tyrosine-p
hosphatase (EC 3.1.3.48), low molecular weight - fission yea
st (Schizosaccharomyces pombe) >gp:gp|L33929|YSPLMPTP_1 Schi
zosaccharomyces pombe low Mr protein tyrosine phosphatase mR
NA, complete cds. NID: g602991.
gtgatactaatgatacatgtagcatttgtatgtctcggtaatatatgtcgttctccaatg
gctgaggctatcatgagacaaagactacaagaaagaggtatttcagatataaaagttcat
tctagaggaacaggacgttggaatttaggcgaacctccacataacggaacacaaaaaatt
ctacagaagtaccatattccttatgatggtatggtgagtgaacttttcgaacctgatgat
gattttgactatattattgctatggaccaaagtaacgtagacaatatcaaacaaatcaat
ccaacttacaaggacaattgttcaaattgctagaatttagtaacatggaagagagtgat
gtaccagatccatactacacaaataattttgaaggtgttttcgagatggtgcaatcatct
tgtgataatttaatagactacatcgtaaaagatgcaaatttgaaagagaggtaa

Sequence 2888
VILMIHVAFVCLGNICRSPMAEAIMRQRLQERGISDIKVHSRGTGRWNLGEPPHNGTQKI
LQKYHIPYDGMVSELFEPDDDFDYIIAMDQSNVDNIKQINPNLQGQLFKLLEFSNMEESD
VPDPYYTNNFEGVFEMVQSSCDNLIDYIVKDANLKER*

Sequence 2889
Contig_0488_pos_5255_6256
>sp:sp|P44770|OTC_HAEIN ORNITHINE CARBAMOYLTRANSFERASE (EC
2.1.3.3) (OTCASE). >pir:pir|H64079|H64079 ornithine carbamoy
ltransferase (arcB) homolog - Haemophilus influenzae (strain
Rd KW20) >gp:gp|U32741|U32741_4 Haemophilus influenzae Rd s
ection 56 of 163 of the complete genome. NID: g1573582.
atgaaaaatttacgtaacagaagcttttttaactttattagactttttcacgacaagaggta
gaatttttattaacactctccgaagatttgaagcgtgccaaatatatcggcactgaaaag
cctatgctaaaaaataaaaatatcgcgcttcttttgaaaaagattccactagaacacgt
tgcgcattcgaagttgccgcacatgatcaaggtgcacacgtcacttatcttggacc:aca
ggttctcaaatgggtaaaaaagaaactgctaaagatacagcacgtgtacttggtggtatg
tatgatggtattgagtaccgaggtttctctcaacgtactgtagaaacattagcgcaatat
tcaggtgttccggtatggaatggattaaccgatgaagatcaccctacacaagtgcttgct
gacttttaactgctaaagaagtattgaaaaaagagtatgctgatatcaactttacttat
gttggcgatggacgtaacaatgttgctaacgcattaatgcaaggtgctgccattatgggt
atgaatttccatcttgtttgtcctaaagaactcaatccgacagaagaattattaaatcgt

tgcgaacgtattgcgacggaaaatggcggtaacattttaataacagatgatattgataaa
ggcgtgaaagattctgatgttatttatacagatgtttgggtatcaatggcgaacctgat
gaagtatggcaagaacgccttaaactttttaaaaccatatcaagttaaccaagcattatta
gaaaaaacaggcaatccaaatgttatttttgaacattgtttaccttcttttccacaatgca
gaaactaaaattggtcaacaaatttatgaaaaatatggcattagtgaaatggaagtcact
gatgatgtcttcgaaagcaaagcttctgtagtattccaagaagctgagaatagaatgcat
acaattaaagcggtcatggtagcaactttaggagaattctaa

Sequence 2890

MKNLRNRSFLTLLDFSRQEVEFLLTLSEDLKRAKYIGTEKPMLKNKNIALLFEKDSTRTR
CAFEVAAHDQGAHVTYLGPTGSQMGKKETAKDTARVLGGMYDGIEYRGFSQRTVETLAQY
SGVPVWNGLTDEDHPTQVLADFLTAKEVLKKEYADINFTYVGDGRNNVANALMQGAAIMG
MNFHLVCPKELNPTEELLNRCERIATENGGNILITDDIDKGVKDSDVIYTDVWVSMGEPD
EVWQERLKLLKPYQVNQALLEKTGNPNVIFEHCLPSFHNAETKIGQQIYEKYGISEMEVT
DDVFESKASVVFQEAENRMHTIKAVMVATLGEF*

Sequence 2891
Contig_0488_pos_6270_7211
is similar to (with p-value 3.0e-76)
  >sp:sp|Q46807|ARCL_ECOLI CARBAMATE KINASE-LIKE PROTEIN 1. >
gp:gp|U28375|ECU28375_24 Escherichia coli K-12 genome; appro
ximately 64 to 65 minutes. NID: g887800. >gp:gp|AE000370|AE0
00370_9 Escherichia coli K-12 MG1655 section 260 of 400 of t
he complete genome. NID: g2367170.
gtgagtgaaatggctaaaattgtagtagctttaggtggaaacgctttaggaaaatcacca
caagaacaacttgaattagtaaaaaatacagctaaatccctagtaggattaattactaaa
ggtcacgaaattgtgattagtcacggtaatggaccacaagtaggaagtattaaccttggt
ctgaattatgcagctgaacacgatcaaggtcctgcttttccatttgctgaatgtggcgct
atgagtcaagcctacatcggctatcaacttcaagaaagtttacaaaatgaacttcattca
atgggcatagataagcaagttgtcacactagttacccaagtagaagttgatgaaggcgat
ccagcttttaatagtccaagtaaacccatcggtctgttctacactaaagaagaagcaaat
cgtattcaacaggaaaaaggttatcaatttgtagaagatgctggtcgaggttaccgtcgc
gttgtaccatcaccacaaccaatatctattatcgaactggaaagtattaaaactctagta
gaaaatgacacactcgtcatcgctgcaggtggaggtggtataccagtcattcgcgaacag
catgatagctttaaaggtatagatgccgtcatcgataaagacaaaacaagtgcattatta
ggtgctgatattcactgtgatcaactcattattttaacagcgattgattatgtttatatc
aactatcatactgaccaacaacaagcacttaaaacaacaaatatagatacgcttaaaaca
tatattgaagaagaacaatttgccaaaggcagcatgctacctaaaatcgaatctgccatc
tcctttattgaaaataatcctaacggtagcgtgctcatcacatcattaaatcaattagat
gcagcactagaaggtaaaattggcacactcattacaaagtaa

Sequence 2892

VSEMAKIVVALGGNALGKSPQEQLELVKNTAKSLVGLITKGHEIVISHGNGPQVGSINLG
LNYAAEHDQGPAFPFAECGAMSQAYIGYQLQESLQNELHSMGIDKQVVTLVTQVEVDEGD
PAFNSPSKPIGLFYTKEEANRIQQEKGYQFVEDAGRGYRRVVPSPQPISIIELESIKTLV
ENDTLVIAAGGGGIPVIREQHDSFKGIDAVIDKDKTSALLGADIHCDQLIILTAIDYVYI
NYHTDQQQALKTTNIDTLKTYIEEEQFAKGSMLPKIESAISFIENNPNGSVLITSLNQLD
AALEGKIGTLITK*

Sequence 2893
Contig_0489_pos_5066_5410
is similar to (with p-value 5.0e-39)
  >sp:sp|P37941|ODBB_BACSU 2-OXOISOVALERATE DEHYDROGENASE BET
A SUBUNIT (EC 1.2.4.4) (BRANCHED- CHAIN ALPHA-KETO ACID DEHY
DROGENASE COMPONENT BETA CHAIN (E1)) (BCKDH E1-BETA). >pir:p
ir|S32487|S32487 3-methyl-2-oxobutanoate dehydrogenase (lipo
amide) (EC 1.2.4.4) E1 beta chain - Bacillus subtilis >gp:gp
|M97391|BACBRANCH_2 Bacillus subtilis branched chain alpha-k
eto acid dehydrogenase E1-alpha, branched chain alpha-keto a

cid dehydrogenase E1-beta, and branched chain alpha-keto aci
d dehydrogenase E2, complete cds. NID: g142610. >gp:gp|D8443
2|BACJH642_247 Bacillus subtilis DNA, 283 Kb region containi
ng skin element. NID: g2627063. >gp:gp|Z99116|BSUB0013_115 B
acillus subtilis complete genome (section 13 of 21): from 23
95261 to 2613730. NID: g2634723.
gtgaattactgtttacaagctgcagatattttggcaaatgacggcatcgatgttgaagta
gtcgacttaagaacagtttatccactagataaagcaactatcattgaacgctctcaacgt
actggtaaagttcttcttgttactgaagataatctagagggaagcattatgtctgaagta
tctgcaattatagctgaaaactgtctgttcgatttagatgcgccaatcatgcgattagct
gcaccggatgtcccatctatgccattttcaccaacattagaaaatgaaattatgatgaac
ccagaaaagatacaggacaaaatgcgtgaactcgcacaattttaa

Sequence 2894
VNYCLQAADILANDGIDVEVVDLRTVYPLDKATIIERSQRTGKVLLVTEDNLEGSIMSEV
SAIIAENCLFDLDAPIMRLAAPDVPSMPFSPTLENEIMMNPEKIQDKMRELAQF*

Sequence 2895
Contig_0493_pos_2490_1000
 >sp:sp|O69282|MQO_CORGL MALATE:QUINONE OXIDOREDUCTASE (EC 1
.1.99.16) (MALATE DEHYDROGENASE (ACCEPTOR)) (MQO).
atgcacatgagtgaagcaaatcataaaaacatcgttgttgtaggtgcaggaattattggt
acgtcagtagcgacaatgctttcaaaagtaagtcctaactggcatatcgatatgtttgaa
agactagaaggcgctggtattgaaagttcaaatgaaaataataatgctgggacaggtcat
gcggcattatgtgaattaaactatacagttgaacaagatgatggttcaattgatgcatct
aaagcgcaagaaattaatgaacaattcgaattatctagacaattctggggtaatttagtt
aaaaatggtgatatttctaatcctgaagaatttattcaaccattacctcatatcagtttc
gttatgggaccaacaaacgttaacttttttaagaaaacgttatgaaacactaagaactctt
ccaatgttcgatacaatcgaatatacagaagacatggaaacaatgagaaaatggatgcca
ttaatgatggaaaatcgtgaaccaggtcatcaaatggcagcaagtaaaattgatgaaggt
acagatgtgaactatggtgcgttaacacgtaagttagcacattacttagaacaaaaatct
aatgtttcattaaaatacaatcatgatgttgtagatttaacacaaagagaagatggcaaa
tgggaagttgtcgttgaaaatagagaaactaaagaaaaagtaactaaaatagcagataaa
gtgtttattggtgctggcggtcactctattccgttattacaaaaatctggcgttaaacaa
agagaacacctaggtggtttcccaatcagtggtcaattcttaagatgtacaaacccagat
attattaaacaacatgcggctaaagtttacagtaaagagcctcaaggtaagccaccaatg
actgtaccacaccttgatacacgttatatcaatggtaaacaaacattattatttggtcca
tatgcgaatatcggccctaaattcttgaaattcggttcaaatctagacttattcgaatca
atcaaaccatataacattactacaatgttggcttcagcagttaaaaatgtaccttttaatt
aaatattcaattgatcaaatgatcaaaactaaagaaggttgtatgaactatttaagaaca
tttattcctgatgctaaagatgaagattgggaactttacactgctggtaaacgtgttcaa
gttattaaagatagtgaacaacacgggaaaggtttcgtagtatttggtactgaagttgtc
aattcagacgacaattctatgattgcattattaggtgaatctccaggggcttcaacatca
ttatcagttgtattagaagtttttagagaaaaacttcgctgatgacaaagaagcatgggaa
cctgttgttaaagaaatggtaccaacatacggtcgttcattaattaatgacgaaaaatta
atgagagaaacacgtcgcgaaacttctaaaaacttacatttaaatagataa

Sequence 2896
MHMSEANHKNIVVVGAGIIGTSVATMLSKVSPNWHIDMFERLEGAGIESSNENNNAGTGH
AALCELNYTVEQDDGSIDASKAQEINEQFELSRQFWGNLVKNGDISNPEEFIQPLPHISF
VMGPTNVNFLRKRYETLRTLPMFDTIEYTEDMETMRKWMPLMMENREPGHQMAASKIDEG
TDVNYGALTRKLAHYLEQKSNVSLKYNHDVVDLTQREDGKWEVVVENRETKEKVTKIADK
VFIGAGGHSIPLLQKSGVKQREHLGGFPISGQFLRCTNPDIIKQHAAKVYSKEPQGKPPM
TVPHLDTRYINGKQTLLFGPYANIGPKFLKFGSNLDLFESIKPYNITTMLASAVKNVPLI
KYSIDQMIKTKEGCMNYLRTFIPDAKDEDWELYTAGKRVQVIKDSEQHGKGFVVFGTEVV
NSIDNSMIALLGESPGASTSLSVVLEVLEKNFADDKEAWEPVVKEMVPTYGRSLINDEKL
MRETRRITSKNLHLNR*

Sequence 2897

Contig_0494_pos_12199_12525
>gp:gp|AB009866|AB009866_15 Bacteriophage phi PVL proviral
DNA, complete sequence. NID: g3341907.
atgaattcagcagtagtagagtcagaactgaattctacaccatcttctttggtattacct
tctaattcagggaatgtgaataatcctttaggtaatcccacatactcgcgtgaaccatct
tccatagttttcgcaaacataacagctacatatggtggtgtatcgttaccaactgacacg
atgccgtcctctgattttctaatccgaacaatgcaactctgtcctctaatggtagtttg
tggaaaccagcttctacttcaattgttccgttagcaactgccatttctgcaacttggtta
tcaccatatgccttctcgatgtcttga

Sequence 2898
MNSAVVESELNSTPSSLVLPSNSGNVNNPLGNPTYSREPSSIVFANITATYGGVSLPTDT
MPSSDFSNPNNATLSSNGSLWKPASTSIVPLATAISATWLSPYAFSMS*

Sequence 2899
Contig_0494_pos_16099_14756
is similar to (with p-value 3.0e-17)
>gp:gp|AB009866|AB009866_13 Bacteriophage phi PVL proviral
DNA, complete sequence. NID: g3341907.
atggctaatttagatgagcgcaaaaaagaaatcgctaatctgatttctaaagcgcaagaa
gcagtcgaaaaaggcgacctcgaaactgctcgtaatttaaaagctgatattgatgctcaa
aagaaagagtacgaagaactcgaacagctttcaaaagaaattgaagcgtcagcacctaaa
caagatgaaccacctaaagatgaaggtgcagaagttgaagataacaaagatggtaattct
ggagaagaatcagagaacaaaccttctgatgatgaaccagaaggaacttcagatgaagaa
aaacctgatgatgcaccaaaaccagatgacaaacctgaagaaacaccagaaacacctact
attgaaaaagtagaagaaccaaccagaagaagaattaaaaaagaaaaagacaaaaaagaa
ggagcgaaacgttctatggctaaattaaaccaaaatccagagacaaacgaagaaattcta
gcacctcaacagtacatgaaatcaaaaggggctaaacgtgacaatgttaaatctgatgac
gttggcgtaactatcccagaggatattaaatatattcctgaaaaagaagttaagacagtc
caagacttatcagaattggtacaaaaaacttcagtatcaactgcaagtgggaaatacccg
atcttaaaacgtgctaacgctaaattcaacactgttgctgaattagagaaaaaccctgag
ttagctcgtccggaattcgaaacaatcaattgggaagtagacacttatcgtggatctatt
ccgatttcacaagaagcattagatgattcagttgctaacttaactgctattgtttctgaa
aatattaacgaacaaaaaatcaacactttaaatgaacgtattggtgaagtttttaaaagca
ttcaatcctactagtgtttctaatgttgacgacttaaaagaaattatcaacgttaaatta
gatcctggttatgaccgccaaattatctgtactcaaagtttctatcaaaaactagataca
ttaaaagatggtaacggtcgttatttactacaagacagtatcatcaacactgcaggtaac
actgtgttaggtatgaatgtaacagttgtgcgtgatgacttgttaggtaaaaatggagat
gcattagcatttattggtgatgtaaaacgcggtgtgttatttgcagaccgtacagacgtt
tctgttcaatggattgaaaatgaaatctacggtaaatacttaatgggtgctttccgtttc
gatgtgaaacaggctgataaaaatgctggtttcttcgtaacatttgaagatgcaacagaa
cctagtggggatctaggagcataa

Sequence 2900
MANLDERKKEIANLISKAQEAVEKGDLETARNLKADIDAQKKEYEELEQLSKEIEASAPK
QDEPPKDEGAEVEDNKDGNSGEESENKPSDDEPEGTSDEEKPDDAPKPDDKPEETPETPT
IEKVEEPTEEELKKEKDKKEGAKRSMAKLNQNPETNEEILAFEQYMKSKGAKRDNVKSDD
VGVTIPEDIKYIPEKEVKTVQDLSELVQKTSVSTASGKYPILKRANAKFNTVAELKKNPE
LARPEFETINWEVDTYRGSIPISQEALDDSVANLTAIVSENINEQKINTLNERIGEVLKA
FNPTSVSNVDDLKEIINVKLDPGYDRQIICTQSFYQKLDTLKDGNGRYLLQDSIINTAGN
TVLGMNVTVVRDDLLGKNGDALAFIGDVKRGVLFADRTDVSVQWIENEIYGKYLMGAFRF
DVKQADKNAGFFVTFEDATEPSGDLGA*

Sequence 2901
Contig_0494_pos_14069_13677
is similar to (with p-value 5.0e-25)
>gp:gp|AB009866|AB009866_13 Bacteriophage phi PVL proviral
DNA, complete sequence. NID: g3341907.
atgaaaggcgataaagaaataattgcctatttagaaacgaaatacggtaaatctgctatg

```
        aaacgcataactgactttgcactaactaaagctggcaataaagttgtaagtattatcaaa
        ggtaatatgaaaagttttgaagatactggagaatcagtagaagaaactacactttcaaag
        ccgatgacgataaaaggagtaaggaccgttaaaattcattggcgaggtcctaaacaacgt
        tatcgtattatccacctaaatgaatatggtcactttgatcgttctggaaagtgggttaat
5       acagctggtaaaggtgttattgaaaacgctatgcgtgaaggcagagaaacgtatttcaga
        acagtaaaagaagaaatgagaaagagggtgtaa
```

Sequence 2902
MKGDKEIIAYLETKYGKSAMKRITDFALTKAGNKVVSIIKGNMKSFEDTGESVEETTLSK
10   PMTIKGVRTVKIHWRGPKQRYRIIHLNEYGHFDRSGKWVNTAGKGVIENAMREGRETYFR
TVKEEMRKRV*

Sequence 2903
Contig_0494_pos_12659_12030
15   No hits found
```
        atggcagagaaaaactatcgttcatttacagggttaacagaattttattataaagttcat
        ggagaaggtggagttcaaaaagttgctgatccagaacgcattaaatatttacaagaaatt
        tcagtatctaaagatcaagacatcgagaaggcatatggtgataaccaagttgcagaaatg
        gcagttgctaacggaacaattgaagtagaagctggtttccacaaactaccattagaggac
20      agagttgcattgttcggattagaaaaatcagaggacggcatcgtgtcagttggtaacgat
        acaccaccatatgtagctgttatgtttgcgaaaactatggaagatggttcacgcgagtat
        gtgggattacctaaaggattattcacattccctgaattagaaggtaataccaaagaagat
        ggtgtagaattcagttctgactctactactgctgaattcatgcaagctaaagttaaaggc
        ttcgaagaagaaaaagcaatgttattaggtcacgatgctaaaggcacaactgttatgaaa
25      gacgctatctgggaagctatcttcggtgaatctgcaccaagcagtgatccaaaagaatct
        agtggaacagaatcagaactaggcgcataa
```

Sequence 2904
MAEKNYRSFTGLTEFYYKVHGEGGVQKVADPERIKYLQEISVSKDQDIEKAYGDNQVAEM
30   AVANGTIEVEAGFHKLPLEDRVALFGLEKSEDGIVSVGNDTPPYVAVMFAKTMEDGSREY
VGLPKGLFTFPELEGNTKEDGVEFSSDSTTAEFMQAKVKGFEEEKAMLLGHDAKGTTVMK
DAIWEAIFGESAPSSDPKESSGTESELGA*

Sequence 2905
35   Contig_0494_pos_11140_6422
is similar to (with p-value 1.0e-33)
>gp:gp|AB009866|AB009866_7 Bacteriophage phi PVL proviral D
NA, complete sequence. NID: g3341907.
```
        atgaaggctatgggcgttcaacgtagtatttcggaaataaaacgtagctttaaaggatta
40      aacgctgacttaaaactatctaacaacaactttaagtattccgaaaaaagtttaaattca
        tataagttaagaactagagaattatcgcaagcagtcaaagaatctaaagctaacgttgca
        gcgttgaaagcaaaataccaagaagcagcaagagcatctggtgtgaatagcaaaaaagcc
        gctcaattaaggcaggaatatagtcgacaagctgacaatctcaactatttacaaaacgaa
        ctcgaccaaacacgtggcaaatacagagaaatgattgcagtaagtaaatcatctgtcggt
45      agacttgggcaagcattttctgaaataggacctaagataagatccataggagattcaatg
        aagtcagtcgggcgtaacatgagtttacacgttactgcaccaattgcagcaggttttggc
        gctgcggtgaagaaaagtatagacttcgacgataccatgcgtaaagtaaaagccacatct
        ggtgctactggagatgaatttaaccagcttagaacaaaagcacttcaaatgggccgagat
        actaaattcacggcctctgaatctgctgaagcaatgaactacatggcgcttgctggttgg
50      gacaccaaagatatgctaaaaggtgttggtggtgtaatggatttagctgctgcatctggt
        gaagatttagcaagcgtatctgatattgtaactgataacctaactgcatttggtatgaaa
        gctaaagatagtacccactttgctgatgtttttggctcaaacgagttcaaaagctaatact
        gatgtacgtggtttaggtgatgcgtttaaatatgctgctccagttgctggtgcgttaggt
        tacacggtagaagatacatcaatagctattggtttgatgtctaatgctgggataaaaggt
55      gaaaaagccggcacagcattaagaacaatgtttaccaacctatctaaaccaacaaaagca
        atgaaagacgaaatggataaactaggaatatctattactgatagcaacggtgaaatgtta
        cctatgagagatgtttagatcaacttagaggtaaaatgggcggtctatctaaagaccaa
        caagcagccgcagctagtacaatatttggtaaagaggccatgagtggtgcattagcagtt
        atcaatgcatcagacgaagattataaaaagctaactaaatccatagacggctctaaaggt
```

```
         gcttcaaaaagaatggctaaagaaatggaaggcggtattggtggcgcaatgcgtaaaatg
         aaatcggcaattgaaagtttagcgatttcattaggtgatgcattagccccaatgttatat
         aaagctgctaaatggatcacatcattagcgaataagttttctaatttacctactggcgtt
         caaaaaacgattgcagttgtaggattactcgccgcagctattggtccactactaatggtc
   5     tttggcgttatggcatcaacaattggcactgctataacagtattaggctctttaatgacg
         agtatgagaacactatcatttttatctaaaaccagtgcagcagcgactggtatttggaat
         ggcgttactgccactgctcgtggtatcgcaaatggttatagatatgcggtggctgcatta
         accacttctcagacaatacaggctatgaaaactaaaattgctgcagctgcaacaacagct
         tggactacagttactaaaggtgcaactttagcaactaaaggcttgggattagcaataaga
   10    tttatgactggaccagtcggtatagttattacagccatcggattattagtagcagggctt
         attcatttatggaaaacaaatagctcgtttagaaatgcagttatcggcatttggaattca
         ataaaaatgctgcaatagctatatttggtttttatcaaaccttatattattaatatttgg
         aacgcaattaaaaactctacaattgccatttggaacgcgattaaaagaagtgctgtaata
         atatggaacgctattaaatttgctgttcaacatcctattcaagcattaaaaaatgtctta
   15    tcagctttatggaatggcatgaaaaatgctgctattaaaatctggaccgccttaaagaac
         ggtgttatagcaattattaaagcatatgttgcgcaagtaaaatttaatatcaaccttatt
         aagcgcatagtagttacgatatttaatgctattaagagtttctctattaaagtttggact
         gcattaaaaaatggtgtgttaggaatcgttcgagctttgcgcaaaggtgttctatctgta
         tttaacgcattaaaaaaaggtattgttgtaatatttaatgctgtaaagaatgccgcagtt
   20    aaaatctggacggctataaaaaaatcagtagtgaataaagcaaaagcattatggtctgga
         gttaaaaatacatggaatgcactcaaaaaaggtacaattggcatatttaaagcagttggc
         agtttcatgagttctaaatggaacagtattaaaaaaggtactgttaataaagcgaaagct
         ctatggtcaggcgtcaaaggtgcttggggatcacttaaaaaaaggtactcataacaccatg
         aatgctgtaggtggcttcatgagcaagaaatggaatggaattaaaagtactactgtatct
   25    atagtaaatggcatgaaatcgaaagttatgggcaccatgaataaaatgagagacggtatc
         aaaacagttaccggtaaaattgggaatcttttttggtggaatggttaaaggcgttaaaaaaa
         ggccttaatggattaatcaaaggtgttaactgggtcgcagataaaattaggtatggataaa
         atacctaagattaaaattatctacaggtactcaatctacgcatacacaaagttatattacc
         aacggtaaaatcaataaaggtacaatggctactgtaggagatagaggtcgaggtaatgga
   30    cctggtggctttagaaatgaaatgatccgttatcctaacggaagaatggcactaacacct
         aacaagatacaacaacattcttacctaaaggctcaagtgtatatagtggcgcgcaaaca
         catgctattttatctaattcaggatatgacactaagaagaaaaaactacctaaatttagt
         aaaggtactaaaaagaaagacggtatattagatgttattagctctggtgtaaaaaatgat
         gttaataaagtaaaagacattggtggtaaagcaagagacataggtggtactacgtttgac
   35    aaagcaaaagacataggtacaaaagcacttgataaagctaaagatgtgtctagcactgtt
         atcaagggtattggagatgttttttgattatgtaggtcatcctatgaaattggtaaataaa
         gtctttgagaaagttggttttaacctagactttatgaaaaatgcaccattaccatttgat
         ttaatgacagctatgattaagaaacttaaaaatggtattaaagacttctttaatgaaggt
         ttagactctgcaggcggtggagatggttcttcgttcactaaattcccaattactacgggg
   40    tattatcctaatggtggtgctcctggttatagtttaatggtggtgctcactttggtatt
         gactatggcgctccatatggtacaactatcaatgctaccaatgatggaaatgtaaaagct
         atccacaacttaggtggaggacttgttgcacgacttttaacaggtcagttcacattgttc
         tttatgcacttatctaaaatattaaaacaaggtaaatcaaagctggagaaccaatggct
         aaaacaggtaattcaggtcaatggactactggtccacacgtacacttccaagttgaaaga
   45    ggtcgccatgatgacatcacaaacagagggacagtaaaccctgctaaatggctcaaaggt
         cacggtggtggaaaagttggtggtagtggttctgtaaacgcacgtagagcaattcaaaga
         gcacaatctatttttaggtggacgttataaatcgtcttatattaccgaacaaatgatgaga
         gttgccaaacgtgagtctaacttccaatcagatgcggttaataactgggacatcaacgca
         caaaaaggaacgccttctaaaggtatgttccaaatgattgaaccatctttt agagcatat
   50    gctaaaccaggacacggaaacatcttaaatccaactgacgaagctatatctgctattgcgt
         tacattccaggtaagtgggttcctattatggggagttggagaagtgcatttaaacttgct
         ggagattatgcttatgctacaggcggggttattaacactgctggattatataatttggca
         gaagatggataccctgagatagtaatccctacagatccaagcagacaatcagatgcgatg
         aaattgttacatcttgctgcaagtaaaattagtggaaataacagaaataaacgacctaac
   55    caattacgtacacctaatgttactagtaatacagttgataatgcagaattactactacaa
         atgatagaaatcaacagaaacaaataaacgtgttaatggaaatagcacgaagtaataaa
         actattgaaaaacaaccgaaaggttttttcagaacgcgatgtaagtcaggcacaaggttca
         aggttaagactcgctgcttatagccagggaggtttataa
```

Sequence 2906
MKDMGVQRSISEIKRSFKGLNADLKLSNNNFKYSEKSLNSYKLRTRELSQAVKESKANVA
ALKAKYQEAARASGVNSKKAAQLRQEYSRQADNLNYLQNELDQTRGKYREMIAVS`.SSVG
RLGQAFSEIGPKIRSIGDSMKSVGRNMSLHVTAPIAAGFGAAVKKSIDFDDTMRKVKATS

5   GATGDEFNQLRTKALQMGRDTKFTASESAEAMNYMALAGWDTKDMLKGVGGVMDLAAASG
EDLASVSDIVTDNLTAFGMKAKDSTHFADVLAQTSSKANTDVRGLGDAFKYAAPVAGALG
YTVEDTSIAIGLMSNAGIKGEKAGTALRTMFTNLSKPTKAMKDEMDKLGISITDSNGEML
PMRDVLDQLRGKMGGLSKDQQAAAASTIFGKEAMSGALAVINASDEDYKKLTKSIDGSKG
ASKRMAKEMEGGIGGAMRKMKSAIESLAISLGDALAPMLYKAAKWITSLANKFSNLPTGV

10   QKTIAVVGLLAAAIGPLLMVFGVMASTIGTAITVLGSLMTSMRTLSFLSKTSAAATGIWN
GVTATARGIANGYRYAVAALTTSQTIQAMKTKIAAAATTAWTTVTKGATLATKGLGLAIR
FMTGPVGIVITAIGLLVAGLIHLWKTNSSFRNAVIGIWNSIKNAAIAIFGFIKPYIINIW
NAIKNSTIAIWNAIKRSAVIIWNAIKFAVQHPIQALKNVLSALWNGMKNAAIKIWTALKN
GVIAIIKAYVAQVKFNINLIKRIVVTIFNAIKSFSIKVWTALKNGVLGIVRALRKGVLSV

15   FNALKKGIVVIFNAVKNAAVKIWTAIKKSVVNKAKALWSGVKNTWNALKKGTIGIFKAVG
SFMSSKWNSIKKGTVNKAKALWSGVKGAWGSLKKGTHNTMNAVGGFMSKKWNGIKSTTVS
IVNGMKSKVMGTMNKMRDGIKTVTGKIGNLFGGMVKGVKKGLNGLIKGVNWVADKLGMDK
IPKIKLSTGTQSTHTQSYITNGKINKGTMATVGDRGRGNGPGGFRNEMIRYPNGRMALTP
NKDTTTFLPKGSSVYSGAQTHAILSNSGYDTKKKKLPKFSKGTKKKDGILDVISSGVKND

20   VNKVKDIGGKARDIGGTTFDKAKDIGTKALDKAKDVSSTVIKGIGDVFDYVGHPMKLVNK
VFEKVGFNLDFMKNAPLPFDLMTAMIKKLKNGIKDFFNEGLDSAGGGDGSSFTKFPITTG
YYPNGGAPGYSFNGGAHFGIDYGAPYGTTINATNDGNVKAIHNLGGGLVARLLTGQFTLF
FMHLSKILKQGKIKAGEPMAKTGNSGQWTTGPHVHFQVERGRHDDITNRGTVNPAKWLKG
HGGGKVGGSGSVNARRAIQRAQSILGGRYKSSYITEQMMRVAKRESNFQSDAVNNWDINA

25   QKGTPSKGMFQMIEPSFRAYAKPGHGNILNPTDEAISAMRYIVGKWVPIMGSWRSAFKRA
GDYAYATGGVINTAGLYNLAEDGYPEIVIPTDPSRQSDAMKLLHLAASKISGNNRNKRPN
QLRTPNVTSNTVDNAELLLQMIENQQKQINVLMEIARSNKTIEKQPKGFSERDVSQAQGS
RLRLAAYSQGGL*

30   Sequence 2907
Contig_0497_pos_4846_3491
is similar to (with p-value 9.0e-50)
>gp:gp|Y10528|PACIOAB_1 P.aeruginosa cioA and cioB genes. N
ID: g2208963.

35   atggattcagtagaaatagctcgattgttgacgggtatgacacttgcagtgcatatcata
tttgcaacaattggtgttggtatgccacttatgtttgcaatagcagagtttattggcata
aagaaaaatgatgcgaattatatcacattggctaaaagatggtcgaaaggctacacgata
accgttgctgttggagtcgttacaggtactattattggtttacaactttcacttgtctgg
ccaacatttatgaaaatgggcggtcatgttatcgcattaccactattcatggaaactttt

40   gcattcttttttgaagcaatcttttttaagtatttatttatacacttgggaaagatttaaa
aataaatggacacattttttcatatctatacccgttattataggaggttcattctcagca
ttctttattacatcagtcaattcatttatgaataccccagctggttttgaaattaaaaat
ggtcgtatggtaaatgttcagccattagaagcaatgtttaattcatcgtttatggttcgt
gctttacatgtagttgcaactgcaggtatgacgatggcgtttatattagcagccatcgca

45   gcgtttaaattattacgtcataatcatacagaagatagaacataccatacaaaagctctt
aatttaagcatgattgttggattcatcaataccgttctttcgatgattgcaggagattta
tctgctaaatttttacacaaagttcaaccagataagcttgcagcatatgaatggcattat
gatacgcaatctcatgcgaatttagttcttttcggtgtattaaatgaaaaaacacatgaa
gtttcaggagcactagagattcccggattacttagtttttttagcagataatagctttaat

50   acaaaagttaaaggtttaaatgaatttcctaaaaatgaattacctcctatgatagtgcat
tactttttcgatttgatggtatcaatgggaatattctgttttatcatttcaggactatat
atgttatttttaattgttaaaaagttacgtaaatatgtaactagtaatatgatgttatac
gctatcttgttaactggtcctgcatcaatgttagcgatagaatttggttggttcttaact
gaaatgggtcgtcaaccgtggattatacgaggatatatgcgcgtatcagaagctgctacc

55   caggcggggggaatcacattagtaacgacgctattcggtttactatatctttttattacta
gttacttcggcttatgtactaattagaatgtttaaaaatcaacctgcttataaagatgta
gaaaaagtgattaaagagagaggtgaaacaaaatga

Sequence 2908

715

MDSVEIARLLTGMTLAVHIIFATIGVGMPLMFAIAEFIGIKKNDANYITLAKRWSKGYTI
TVAVGVVTGTIIGLQLSLVWPTFMKMGGHVIALPLFMETFAFFFEAIFLSIYLYTWERFK
NKWTHFFISIPVIIGGSFSAFFITSVNSFMNTPAGFEIKNGRMVNVQPLEAMFNSSFMVR
ALHVVATAGMTMAFILAAIAAFKLLRHNHTEDRTYHTKALNLSMIVGFINTVLSMIAGDL

5   SAKFLHKVQPDKLAAYEWHYDTQSHANLVLFGVLNEKTHEVSGALEIPGLLSFLADNSFN
TKVKGLNEFPKNELPPMIVHYFFDLMVSMGIFCFIISGLYMLFLIVKKLRKYVTSNMMLY
AILLTGPASMLAIEFGWFLTEMGRQPWIIRGYMRVSEAATQAGGITLVTTLFGLLYLLLL
VTSAYVLIRMFKNQPAYKDVEKVIKERGETK*

10   Sequence 2909
Contig_0499_pos_7860_8597
is similar to (with p-value 4.0e-71)
  >sp:sp|Q46807|ARCL_ECOLI CARBAMATE KINASE-LIKE PROTEIN 1. >
gp:gp|U28375|ECU28375_24 Escherichia coli K-12 genome; appro

15   ximately 64 to 65 minutes. NID: g887800. >gp:gp|AE000370|AE0
00370_9 Escherichia coli K-12 MG1655 section 260 of 400 of t
he complete genome. NID: g2367170.
gtgaatcatttgaatttattttttgggaaaaaagcacattcttaaagatattacattctca
ttacctataaatggtgaaatcattggtattgtaggacctaatggtgccggtaaatcatct

20   cttctcaaagcatttatcggagaatttaaagcgaccggtaaacatacattatatgatcac
cctatacatacacaacttagatatattacatatatcccacaaaaagctcatattgattta
gattttcctataaaagtggatcaagtgatactatcaggatgttatgaagacattggatgg
tttaaaaaagcaagtgtagttgaaaaaacgaaacttaaccaattacttaatgatttagaa
ctcgaccatatacgccatcgacaaattgcagaattaagtggtggacaattgcaacgtgtt

25   cttgttgcaagagcacttatgagtaatagtgatatttattgtttagacgaaccttttgtt
ggcatcgacatttatagcgaacaacttattatgaagaaaattaaacatttaagacatatg
ggtaaattgattttaattgttcatcatgacttatcaaaagcagatcaatattttgaccgc
attttattattaaatcaatcgttgcagttttttagggccaacaaaagaagcattatcatct
gaacgattaaatgcaactttcattaattacaaagatgattcgcttttaacactatcctca

30   caagggagtacgaattag

Sequence 2910
VNHLNLFLGKKHILKDITFSLPINGEIIGIVGPNGAGKSSLLKAFIGEFKATGKHTLYDH
PIHTQLRYITYIPQKAHIDLDFPIKVDQVILSGCYEDIGWFKKASVVEKTKLNQLLNDLE

35   LDHIRHRQIAELSGGQLQRVLVARALMSNSDIYCLDEPFVGIDIYSEQLIMKKIKHLRHM
GKLILIVHHDLSKADQYFDRILLLNQSLQFLGPTKEALSSERLNATFINYKDDSLLTLSS
QGSTN*

Sequence 2911
40   Contig_0499_pos_8598_9434
  >gp:gp|X99127|SEABCTS_1 S.epidermidis gene encoding ABC tra
nsport system. NID: g1617427.
atgttagatttcattaaccatttgcttagttatcaatttttaaatcgtgcattaatcaca
tctattttagttggggattgtatgtggaacgatgggtagcattattgttttacgtggtctt

45   tctttaatgggtgatgccatgagtcatgctgttttaccaggtgttgctttatctttctta
tttaatattccaatgtttatcggggcacttgtaacgggaatgcttgcaagtttgtttatt
ggttttattacttcaaacagtaaaacaaaaccagatgctgcaataggaataagtttcact
gcattcctagcatctggcgtcataattattagtttaatcaatagtacaacagatttatat
cacattttatttggcaatttattagcaattacacatcaatcattttggacaacaattgtc

50   attactgtactggttattttacttattattatcttttatagacctttaatgatttcaaca
tttgatgcaacttttagtcgtatgagcgggctgaacacaacattaattcactactttgtc
atgttattactcgcacttgtaactgttgcgagcatacaaacagttggaattatcctttgta
gttgcttactaatcactccagcttctacagcttttttaatcagtaaacaactttacgcc
atgatggtaattgcaagcataatcagcgtgataagttcgattatcggtctatattttagt

55   tatatatataatattccaagtggagcaactattgtaatctgtacctttatgatttatatt
gtaacgctatcaattactagaattaaaaataaacaaaaaaggagcgctttaacgtga

Sequence 2912
MLDFINHLLSYQFLNRALITSILVGIVCGTMGSIIVLRGLSLMGDAMSHAVLPGVALSFL

716

FNIPMFIGALVTGMLASLFIGFITSNSKTKPDAAIGISFTAFLASGVIIISLINSTTDLY
HILFGNLLAITHQSFWTTIVITVLVILLIIIFYRPLMISTFDATFSRMSGLNTTLIHYFV
MLLLALVTVASIQTVGIILVVALLITPASTAFLISKQLYAMMVIASIISVISSIIGLYFS
YIYNIPSGATIVICTFMIYIVTLSITRIKNKQKRSALT*

Sequence 2913
Contig_0499_pos_0_864
 >gp:gp|X99127|SEABCTS_2 S.epidermidis gene encoding ABC tra
nsport system. NID: g1617427.
gtgagtgaaatggctaaaattgtagtagctttaggtggaaacgctttaggaaaatcacca
caagaacaacttgaattagtaaaaaatacagctaaatccctagtaggattaattactaaa
ggtcacgaaattgtgattagtcacggtaatggaccacaagtaggaagtattaaccttggt
ctgaattatgcagctgaacacgatcaaggtcctgcttttccatttgctgaatgtggcgct
atgagtcaagcctacatcggctatcaacttcaagaaagtttacaaaatgaacttcattca
atgggcatagataagcaagttgtcacactagttacccaagtagaagttgatgaaggcgat
ccagcttttaatagtccaagtaaacccatcggtctgttctacactaaagaagaagcaaat
cgtattcaacaggaaaaaggttatcaatttgtagaagatgctggtcgaggttaccgtcgc
gttgtaccatcaccacaaccaatatctattatcgaactggaaagtattaaaactctagta
gaaaatgacacactcgtcatcgctgcaggtggaggtggtataccagtcattcgcgaacag
catgatagcttttaaaggtatagatgccgtcatcgataaagacaaaacaagtgcattatta
ggtgctgatattcactgtgatcaactcattattttaacagcgattgattatgtttatatc
aactatcatactgaccaacaacaagcacttaaaacaacaaatatagatacgcttaaaaca
tatattgaagaagaacaatttgccaaaggcagcatgctacctaaaatcgaatctgccatc
tcctttattgaaaataatcctaac

Sequence 2914
VSEMAKIVVALGGNALGKSPQEQLELVKNTAKSLVGLITKGHEIVISHGNGPQVGSINLG
LNYAAEHDQGPAFPPFAECGAMSQAYIGYQLQESLQNELHSMGIDKQVVTLVTQVEVDEGD
PAFNSPSKPIGLFYTKEEANRIQQEKGYQFVEDAGRGYRRVVPSPQPISIIELESIKTLV
ENDTLVIAAGGGGIPVIREQHDSFKGIDAVIDKDKTSALLGADIHCDQLIILTAIDYVYI
NYHTDQQQALKTTNIDTLKTYIEEEQFAKGSMLPKIESAISFIENNPN

Sequence 2915
Contig_0500_pos_3850_3113
is similar to (with p-value 2.0e-97)
 >sp:sp|P36839|ARGD_BACSU ACETYLORNITHINE AMINOTRANSFERASE (
EC 2.6.1.11) (ACOAT). >pir:pir|S38431|S38431 N-acetylornithi
ne aminotransferase - Bacillus subtilis >gp:gp|Z26919|BSCITB
O_4 B.subtilis (168) DNA for argC-F citrulline biosynthetic
operon. NID: g408113. >gp:gp|Z99109|BSUB0006_199 Bacillus su
btilis complete genome (section 6 of 21): from 999501 to 120
9940. NID: g2633260. >gp:gp|Z99110|BSUB0007_4 Bacillus subti
lis complete genome (section 7 of 21): from 1194391 to 14111
40. NID: g2633472.
atgaaaaacatcatcgtaattaagctcggtggtatagctatagaaaatttaaacgacgca
tttatacaacaaattaatgcttggcaccttgaaaacaaaaaaataattattgttcacggt
ggcggccaagtcatcagtaatttattaactaaaaacaatcattcaactattaaaattgat
ggcatgagagtaactgctaaaaatgatttacctatcatatatgatgctttaattaacata
gttggccatcaacttttagaaagacttaaagaatctaatttagaattttttcaatttaaa
gaaaagataaaagaacttgtaagcgccgaatttttaaataaaaatatctatggttacgtt
ggtaaagttaaagaaatcaacacgatgctattagaaaaaatgttatcacgcgacataata
ccaattatcactagtttgggtgtaaatgagcaaggggagtatcttaatgttaatgctgat
catctcgccaccgcaattgctaagaaactaaaagtcgaaaaattagtatatatgactgat
gtccctggtgtaattgaaaaagataaaacgcttgctactcttacaattaatgaagcaaaa
acaaaaattgaaaataaaaataattacaggcggaatgatacctaaaattgagagtgcaatc
caaacattagaatctggtgttgaatcgattttaattgcaaataatttacaaaaaggaaca
atcataagggggtgattaa

Sequence 2916

MKNIIVIKLGGIAIENLNDAFIQQINAWHLENKKIIIVHGGGQVISNLLTKNNHSTIKID
GMRVTAKNDLPIIYDALINIVGHQLLERLKESNLEFFQFKEKIKELVSAEFLNKNIYGYV
GKVKEINTMLLEKMLSRDIIPIITSLGVNEQGEYLNVNADHLATAIAKKLKVEKLVYMTD
VPGVIEKDKTLATLTINEAKTKIENKIITGGMIPKIESAIQTLESGVESILIANNLQKGT
5   IIRGD*

Sequence 2917
Contig_0500_pos_3110_1983
is similar to (with p-value 1.0e-34)
10   >sp:sp|P36840|ARGB_BACSU ACETYLGLUTAMATE KINASE (EC 2.7.2.8
) (NAG KINASE) (AGK) (N-ACETYLGLUTAMATE 5-PHOSPHOTRANSFERASE
). >pir:pir|S38430|S38430 N-acetylglutamate 5-phosphotransfe
rase - Bacillus subtilis >gp:gp|Z26919|BSCITBO_3 B.subtilis
(168) DNA for argC-F citrulline biosynthetic operon. NID: g4
15   08113. >gp:gp|Z99109|BSUB0006_198 Bacillus subtilis complete
genome (section 6 of 21): from 999501 to 1209940. NID: g263
3260. >gp:gp|Z99110|BSUB0007_3 Bacillus subtilis complete ge
nome (section 7 of 21): from 1194391 to 1411140. NID: g26334
72.
20   atgagttatctttttaataattacaagcgtgacaatatagagtttgttgatgctaatcaa
aatgaattaattgataaagataataatgtctacctagatttttcgtcaggtataggtgtg
acaaatctggggttttaatatggaaatttaccaagcagtttataatcaactgaatttaata
tggcattcacccaatttatacctaagtagtatccaagaggaagtggctcaaaaattaatt
ggtcaacgagattatttagctttctttttgtaatagcggaacagaagcgaatgaggcagct
25   atcaaactcgcacgtaaagctactggtaagtcggaaattattgcttttaaaaagtctttt
cacggcagaacgtacggcgcaatgtctgcaacaggacagaagaaaattacagatcaattt
ggtccggttgttcctggattcaaatttgctattttttaatgattttaattcatttaaatca
ttaacttcaaataatactgctgctgtaattatagaaataattcaaggtgaatcaggagta
ctacctgctgattctttatttatgaagcaattaaatgagtattgtaaacaaaaagatatc
30   cttataattgtagacgaggttcaaacgggcataggtagaaccggtaagttatatgctcat
gaacattatcaattgtctccagatatcatcacattagctaaaggattaggtaatggcctt
cctattggagcaatgttaggcaaaaaaaatttaggtcatgcatttggctacggttctcat
ggtacaacattcggtggaaatagattatcattggctgctgcaaaccaaacgctttctatc
attaatgatgctgatttgctgaatgatgttcaatcaaaggggcaatttcttattgaaaac
35   ttaagaaaaagtttagtaaataaaagaaatgtaattgaagtacgtggtgtaggtttaatg
gtaggaatagaggtcactaatgatcctagtcaagtagtgcgagaagctaaacgtatggggg
ttaatcattttaacagctggtaaaaatgtgattaggttattaccgccattgaccatcact
aaaaaacaattagaaaaaggtatagaaatattaactgaaatcatttga

40   Sequence 2918
MSYLFNNYKRDNIEFVDANQNELIDKDNNVYLDFSSGIGVTNLGFNMEIYQAVYNQLNLI
WHSPNLYLSSIQEEVAQKLIGQRDYLAFFCNSGTEANEAAIKLARKATGKSEIIAFKKSF
HGRTYGAMSATGQKKITDQFGPVVPGFKFAIFNDFNSFKSLTSNNTAAVIIEIIQGESGV
LPADSLFMKQLNEYCKQKDILIIVDEVQTGIGRTGKLYAHEHYQLSPDIITLAKGLGNGL
45   PIGAMLGKKNLGHAFGYGSHGTTFGGNRLSLAAANQTLSIINDADLLNDVQSKGQFLIEN
LRKSLVNKRNVIEVRGVGLMVGIEVTNDPSQVVREAKRMGLIILTAGKNVIRLLPPLTIT
KKQLEKGIEILTEII*

Sequence 2919
50   Contig_0502_pos_13693_14094
is similar to (with p-value 7.0e-44)
>sp:sp|P22806|BIOF_BACSH 8-AMINO-7-OXONONANOATE SYNTHASE (E
C 2.3.1.47) (7-KETO-8-AMINO- PELARGONIC ACID SYNTHETASE) (7-
KAP SYNTHETASE) (L-ALANINE--PIMELYL COA LIGASE). >pir:pir|JQ
55   0512|JQ0512 8-amino-7-oxononanoate synthase (EC 2.3.1.47) -
Bacillus sphaericus >gp:gp|M29291|BACBIOXWF_3 B.sphaericus b
ioXWF operon genes, complete cds. NID: g142592.
gtgcacctgattgtgctggcgtacctggttccactggtttacctggttctgctggcgtac
ctggttccgctggtttacctggttccgctggtttacctggttccgctggtttacctggtt

ccgctggtttacctggttctgctggtttacctggttccgctggtttacctggttctgctg
gcgtacctggttccgctggtttacctggttccgctggtttacctggttctgctggtttac
ctggttctgctggcgtacctggttccgctggtttacctggttctgctggcgtacctggtt
ccgctggtttacctggttccgctggtttacctggttccgctggtttacctggttctgctt
5   ttgttggaccatactcaacaatttcgtcaacaggttgtttag

Sequence 2920
VHLIVLAYLVPLVYLVLLAYLVPLVYLVPLVYLVPLVYLVPLVYLVLLVYLVPLVYLVLL
AYLVPLVYLVPLVYLVLLVYLVLLAYLVPLVYLVLLAYLVPLVYLVPLVYLVPLVYLVLL
10   LLDHTQQFRQQVV*

Sequence 2921
Contig_0502_pos_15619_13529
is similar to (with p-value 3.0e-21)
15   >sp:sp|P22818|BIOD_BACSH DETHIOBIOTIN SYNTHETASE (EC 6.3.3.
3) (DETHIOBIOTIN SYNTHASE) (DTB SYNTHETASE) (DTBS). >pir:pir
|JQ0506|JQ0506 dethiobiotin synthase (EC 6.3.3.3) - Bacillus
sphaericus >gp:gp|M29292|BACBIODAYB_1 Bacillus sphaericus I
F03525 bioDAYB operon encoding dethiobiotin synthase (bioD),
20   adenosylmethionine-8-amino-7-oxononanoate aminotransferase
(bioA), biotin synthase (bioY) and bioB genes, complete cds.
NID: g142587.
gtggatgagatcgttcattatggtggcgaagaaatcaagccaggccataaggatgaattt
gatccaaacgcaccgaaaggtagccaagaggacgttccaggtaaaccaggagttaaaaat
25   cctgatacaggcgaagtagtcacaccaccagtggatgatgtgacaaaatatggtccagtt
gatggagattcgattacgtcaacggaagaaattccattcgacaagaaacgtgaattcaat
cctgatttaaaaccaggtgaagagcgtgttaaacagaaaggtgaaccaggaacaaaaaca
attacaaccaccaacaactaagaacccattaacaggggaaaaagttggcgaaggtgacca
acagaaaaaataacaaaacaaccagtagatgaaatcacagaatatggtggcgaagaaatc
30   aagccaggccataaagatgaatttgatccaaatgcaccgaaaggtagccaagaggacgtt
ccaggtaagccaggagttaaaaatcctgatacaggcgaagtagtcacaccaccagtggat
gatgtgacaaaatatggtccagttgatggagatccgattacgtcaacggaagaaattccg
tttgataaaaaacgcgaatttaatccaaacttagcgccaggtacagagaaagtcgttcaa
aaaggtgaaccaggaacaaaaacaattacaacaccaacaactaagaacccattaacaggg
35   gaaaaagttggtgaaggtgaaccaacagaaaaaataacaaaacaaccagtggatgagatc
gttcattatggtggcgaagaaatcaagccaggccataaggatgaatttgatccaaacgca
ccgaaaggtagccaagaggacgttccaggtaaaccaggagttaaaaatcctgatacaggc
gaagtagtcacaccaccagtggatgatgtgacaaaatatggtccagttgatggagattcg
attacgtcaacggaagaaattccgtttgataaaaaacgcgaatttgatccaaacttagcg
40   ccaggtacagagaaagtcgttcaaaaaggtgaaccaggaacaaaaacaattacaacgcca
acaactaagaacccattaacaggagaaaaagttggcgaaggtgaaccaacagaaaaaata
acaaaacaaccagtggatgagattgttcattatggtggtgaacaaataccacaaggtcat
aaagatgaatttgatccaaatgcacctgtagatagtaaaactgaagttccaggtaaacca
ggagttaaaaatcctgatacaggtgaagttgttaccccaccagtggatgatgtgacaaaa
45   tatggtccgaaagttggtaatccaatcacatcaacggaagagattccatttgataagaaa
cgtgtatttaatcctgatttaaaaccaggtgaagagcgcgttaaacaaaaggtgaacca
gg ?acaaaaacaattacaacaccaatattagttaatcctattacaggagaaaaagttggc
gaaggtaaatcaacagaaaaagtcactaaacaacctgttgacgaaattgttgagtatggt
ccaacaaaagcagaaccaggtaaaccagcggaaccaggtaaaccagcggaaccaggtaaa
50   ccagcggaaccaggtacgccagcagaaccaggtaaaccagcggaaccaggtacgccagca
gaaccaggtaaaccagcagaaccaggtaaaccagcggaaccaggtaaaccagcggaacca
ggtacgccagcagaaccaggtaaaccagcggaaccaggtaaaccagcagaaccaggtaaa
ccagcggaaccaggtaaaccagcggaaccaggtaaaccagcggaaccaggtaaaccagcg
gaaccaggtacgccagcagaaccaggtaaaccagtggaaccaggtacgccagcacaatca
55   ggtgcaccagaacaaccaaatagatcaatgcattcaacagataataaaaatcaattacct
gatacaggtgaaaatcgtcaagctaatgagggaactttagtcggatctctattagcaatt
gtcggatcattgttcatatttggtcgtcgtaaaaaaggtaatgaaaaataa

Sequence 2922

```
VDEIVHYGGEEIKPGHKDEFDPNAPKGSQEDVPGKPGVKNPDTGEVVTPPVDDVTKYGPV
DGDSITSTEEIPFDKKREFNPDLKPGEERVKQKGEPGTKTITTPTTKNPLTGEKVGEGEP
TEKITKQPVDEITEYGGEEIKPGHKDEFDPNAPKGSQEDVPGKPGVKNPDTGEVVTPPVD
DVTKYGPVDGDPITSTEEIPFDKKREFNPNLAPGTEKVVQKGEPGTKTITTPTTKNPLTG
EKVGEGEPTEKITKQPVDEIVHYGGEEIKPGHKDEFDPNAPKGSQEDVPGKPGVKNPDTG
EVVTPPVDDVTKYGPVDGDSITSTEEIPFDKKREFDPNLAPGTEKVVQKGEPGTKTITTP
TTKNPLTGEKVGEGEPTEKITKQPVDEIVHYGGEQIPQGHKDEFDPNAPVDSKTEVPGKP
GVKNPDTGEVVTPPVDDVTKYGPKVGNPITSTEEIPFDKKRVFNPDLKPGEERVKQKGEP
GTKTITTPILVNPITGEKVGEGKSTEKVTKQPVDEIVEYGPTKAEPGKPAEPGKPAEPGK
PAEPGTPAEPGKPAEPGTPAEPGKPAEPGKPAEPGKPAEPGTPAEPGKPAEPGKPAEPGK
PAEPGKPAEPGKPAEPGKPAEPGTPAEPGKPVEPGTPAQSGAPEQPNRSMHSTDNKNQLP
DTGENRQANEGTLVGSLLAIVGSLFIFGRRKKGNEK*
```

Sequence 2923
Contig_0502_pos_12781_12110
is similar to (with p-value 2.0e-34)
  >gp:gp|U67763|PBU67763_1 Plasmodium berghei thrombospondin
related adhesion protein (PbTRAP) gene, complete cds. NID: g
1906578.

```
atgaatatatttgtaactggaacgaatactgacatcggtaaaacgtatgtcactaaatat
ctctataaggcattaagaacgaggggatatcgggtatgtattttcaaaccttttcaaact
gaagaaattggtggaggtagatacccagatttagaaatttataaaaacgaatgcgattta
gactatgacgttacgtctctttacacattcaaagatccagtttcaccacatttagcattc
aaaattgaaaggcatcagcaattgaaccatcaaacaatgatagacaaactcgaatcacta
gaagcacaattcgatatgattctcattgaaggtgcaggtggtattgcagtgcctatctat
gaatacagtgaccattttatatgacaacagatttaattaaagacacatcggatttcatt
gtgagtgtcttaccttcaaagttaggtgcgattaatgatgccatcgttcaccagaaatat
attgatcatcaagaacttcccccgaatgtgttaatcatgaataactatactgatagtgct
attgaacaggataatttacataccattgaaaaattaatacacaagtctgtttatacgttg
ggtcatcaagcgactcaagaaagcttttccgaagcatttatacaacgaattataggagga
tccaatggctaa
```

Sequence 2924
MNIFVTGTNTDIGKTYVTKYLYKALRTRGYRVCIFKPFQTEEIGGGRYPDLEIYKNECDL
DYDVTSLYTFKDPVSPHLAFKIERHQQLNHQTMIDKLESLEAQFDMILIEGAGGIAVPIY
EYSDHFYMTTDLIKDTSDFIVSVLPSKLGAINDAIVHQKYIDHQELPPNVLIMNNYTDSA
IEQDNLHTIEKLIHKSVYTLGHQATQESFSEAFIQRIIGGSNG*

Sequence 2925
Contig_0502_pos_10751_9618
is similar to (with p-value 8.0e-36)
  >gp:gp|U67763|PBU67763_1 Plasmodium berghei thrombospondin
related adhesion protein (PbTRAP) gene, complete cds. NID: g
1906578.

```
atggacattaaagcacagttaaaacagattcaagataaaggtttatatagagagcttcag
ccgattcagtccgtagaaaaacaatatatttatatcaatgaccaatcttatattaatttt
acttcgaacgattatctcggtataggacaagttgaatatcaacctcaaaatttcttagat
tttataaagacatatagtatccatctatcaagttctagattagtgagtggcaattcagtt
gtttatcagcaattagaacaggaaattagcgagcattttaattttgaagacgccttaatt
tttaatagtggttacgatgcgaatttggcggtatttaatatttttaaaaataataatatt
gttattttttctgatcaacagaatcatgccagtataatagacggtattaaattaagtggt
ttatcaaaagtgatttatcaacatttaaactatgatgacctggaaagtcatttagcacgg
cacaccaatccagatgttcaaaaagtaattgtctctgatagtgtgttttctactaatggc
actaaagcagatattaataggctagtacatctcaagcaacgttacaatgcgattttaatt
attgacgcatctcatagtttaggattaaatctctttgagtatcatgcagacattgacata
gttacttcaagtttatctaaagcgtggggagcccatggtggcgtcatattcagttcaaaa
gatataaaagatttaatcattaataaaggtcgttcgcttatctactcgagtagtttacct
agctatcatttgtattttattcaagtgagcttacaacatgtgattgaagatacatacaga
cgagagaagttgaatgtacttagtgaatattttaatcaccaattcatggaattatttccc
```

```
gatcaaccattatctaatacacctatcaaaaatatcgtttgtgatagtttggcttcagca
caagcacaatacgacatgttatttgaacatggtatatttgtcagttatttaaggtatcca
acagtgtcacagctaacattaagaatttcattatcctattttcatgacacagatgatatt
gatcgacttttcaatgtaatgaaacaatacgatgaaggtgatagctatgtatag
```

5

Sequence 2926
MDIKAQLKQIQDKGLYRELQPIQSVEKQYIYINDQSYINFTSNDYLGIGQVEYQPQNFLD
FIKTYSIHLSSSRLVSGNSVVYQQLEQEISEHFNFEDALIFNSGYDANLAVFNIFKNNNI
VIFSDQQNHASIIDGIKLSGLSKVIYQHLNYDDLESHLARHTNPDVQKVIVSDSVFSTNG

10   TKADINRLVHLKQRYNAILIIDASHSLGLNLFEYHADIDIVTSSLSKAWGAHGGVIFSSK
DIKDLIINKGRSLIYSSSLPSYHLYFIQVSLQHVIEDTYRREKLNVLSEYFNHQFMELFP
DQPLSNTPIKNIVCDSLASAQAQYDMLFEHGIFVSYLRYPTVSQLTLRISLSYFHDTDDI
DRLFNVMKQYDEGDSYV*

15   Sequence 2927
Contig_0509_pos_6421_7818
    >sp:sp|P13408|UHPT_ECOLI HEXOSE PHOSPHATE TRANSPORT PROTEIN
    . >pir:pir|A30395|MMECHP hexose phosphate transport protein
    uhpT - Escherichia coli >gp:gp|M17102|ECOUHP_5 E.coli uhp op
20   eron encoding UhpA, UhpB, UhpC, and UhpT protein, (encoding
    hexose phosphate transport protein), complete cds, and an il
    vBN operon encoded protein, 3' end. NID: g148110. >gp:gp|M89
    479|ECOUHPABCT_4 Escherichia coli uhpABCT operon encoding he
    xosephosphate utilization protein (uhpA) gene, complete cds,
25   and hexosephosphate transport protein (uhpB, uhpC, uhpT) ge
    nes, complete cds. NID: g148116. >gp:gp|AE000444|AE000444_5
    Escherichia coli K-12 MG1655 section 334 of 400 of the compl
    ete genome. NID: g2367258.
    atgatagggggtgttagggatgaactttttgatattcataaaatgccaaacaaagggata
30   ccattagctgtacaacgcaaattatggctcagaaactttatgcaagcgttttttgtcgta
    ttctttgtttacatggcgatgtatttaattcgaaacaatttttaaagcggcacaacagtta
    ttaaaagaagaaatcggattaacaacattagaactaggttatataggattagcgtttagt
    attacttacggtttaggaaaaacaatactcggttatttcgttgatgggcgtaatacgaaa
    cgtattatttccttcttattaatattatctgcgattacagtacttattatgggatttgta
35   ttaagttatttcggttctgtgatggggctattaattgtattgtgggggcttaacggtata
    tttcaatctgtgggtgggcctgcaagttactcaacgatttcaaggtgggcgcctcgaaca
    ·aagcgcggtcgttatttaggcttttggaatacatcacataacattggtggtgctattgct
    ggtggtgtcgcactttggggcgcgaatacattttttccacggtaatgtggttggaatgttt
    attttttccttccgtcatcgctttaatcattgggattgtgacattattttattggtaaagat
40   gatccagaggaattaggttggaatcgtgccgaagaaatttgggaagagcctatcgaccaa
    gaaaacattgattctcaaggtatgactaaatgggatatctttaaaaaatatatccttgga
    aatcctgtgatttggattttgtgtatctctaatgttttgtatatatcgtgcgtattggt
    attgataactgggcaccgctatacgtatcagagcatttacattttaataaaggtgatgcg
    gtgaatactattttttactttgaaataggtgcattagtagctagtttattgtggggctat
45   atctcagatttattaaaaggtcgtcgtgcgattgtagcgattggatgtatgtttatgatc
    acctttgttgtactcttttataccaatgcaacaagcgtgacaatggtcaatatttctcta
    tttgcattaggcgctttaatcttcggtccacagttactcattggtgtatctctgactggc
    tttgttcctaaaaatgcaattagtgtcgctaacggtatgacaggttcatttgcatatcta
    ttcggggattcaatggctaaagtgggtctggctgcaatcgctgatccaacacgtaatggt
50   ttaaatattttgggtatacgttgagtggttggacagatgtctttattgtattctatgta
    gctttattcttaggaatgatattattagccattgttgcttattacgaagaaaagaaaatt
    agaaaattaaaaatttaa

Sequence 2928
55   MIGVLGMNFFDIHKMPNKGIPLAVQRKLWLRNFMQAFFVVFFVYMAMYLIRNNFKAAQPL
LKEEIGLTTLELGYIGLAFSITYGLGKTILGYFVDGRNTKRIISFLLILSAITVLIMGFV
LSYFGSVMGLLIVLWGLNGIFQSVGGPASYSTISRWAPRTKRGRYLGFWNTSHNIGGAIA
GGVALWGANTFFHGNVVGMFIFPSVIALIIGIVTLFIGKDDPEELGWNRAEEIWEEPIDQ
ENIDSQGMTKWDIFKKYILGNPVIWILCISNVFVYIVRIGIDNWAPLYVSEHLHFNKGDA
```

VNTIFYFEIGALVASLLWGYISDLLKGRRAIVAIGCMFMITFVVLFYTNATSVTMVNISL
FALGALIFGPQLLIGVSLTGFVPKNAISVANGMTGSFAYLFGDSMAKVGLAAIADFTRNG
LNIFGYTLSGWTDVFIVFYVALFLGMILLAIVAYYEEKKIRKLKI*


5    Sequence 2929
     Contig_0517_pos_10178_9693
     is similar to (with p-value 2.0e-26)
      >sp:sp|P55978|GREA_HELPY TRANSCRIPTION ELONGATION FACTOR GR
     EA (TRANSCRIPT CLEAVAGE FACTOR GREA). >gp:gp|AE000596|HPAE00
10   0596_20 Helicobacter pylori section 74 of 134 of the complet
     e genome. NID: g2313982.
     atgagatttatggaaaaccaaaaacaatatcctatgactcaagaaggttatgagaaactt
     gaacaagaattagaagaattaaaaacggttaaaagacctgaggtagttgaaaaaataaaa
     gtagctcgttcatttggagacctatctgagaactctgaatatgatgctgctaaagatgaa
15   caaggctttattgaacaagatatacaacgtattgaacatatgattagaaatgcgttaatc
     attgaagataacggtgataacaatgtagttcaaattggtaaaacagttacttttattgaa
     ttacctggagatgaagaagaaagttatcaaatcgttggttctgctgaagctgacgcattt
     aaaggaaaaatttctaacgaatctccaatggcaaaagcactaatcggtaaaggattaaat
     gatcaagtacgtgttccacttcctaacggtggcgaaatgaatgttaaaatcgttgaaatt
20   aaataa


     Sequence 2930
     MRFMENCKQYPMTQEGYEKLEQELEELKTVKRPEVVEKIKVARSFGDLSENSEYDAAKDE
     QGFIEQDIQRIEHMIRNALIIEDNGDNNVVQIGKTVTFIELPGDEEESYQIVGSAEADAF
25   KGKISNESPMAKALIGKGLNDQVRVPLPNGGEMNVKIVEIK*


     Sequence 2931
     Contig_0517_pos_9512_8913
     is similar to (with p-value 1.0e-57)
30    >sp:sp|P24247|PFS_ECOLI PFS PROTEIN (P46). >pir:pir|S45227|
     S45227 purine nucleoside phosphorylase homolog - Escherichia
      coli >gp:gp|D26562|ECO82K_47 Escherichia coli genome, 2.4-4
     .1 min region (110,917-193,643 bp from 0 min). NID: g473770.
      >gp:gp|U70214|ECU70214_10 Escherichia coli chromosome minut
35   es 4-6. NID: g1552727. >gp:gp|AE000125|AE000125_6 Escherichi
     a coli K-12 MG1655 section 15 of 400 of the complete genome.
      NID: g1786348. >gp:gp|U24438|ECU24438_1 Escherichia coli MT
     A/SAH nucleosidase gene, complete cds. NID: g2981266.
     atgcgctatgcgtatggttcagaaaaggttttaccgttatctaaaaaaaacatctcatcat
40   gctagaatatatattggtcagccaaccaaaatattcaacacgaggagatgctttttttatg
     tcatctgacacaaacagtttagcacatacaaaatggaattgtaagtatcacatagtcttt
     gtaccaaaatatagaagacaagtgatatacggaaaaatcaaaagagatattggagttatt
     ttacgtcaactatgtgaaagaaaaggcgtagaaataatagaagcagaagcatctaaagat
     catattcatatgttagttagtattccacctaaattaggagtatcctcatttgttggatat
45   ttaaaaggtaaaagcagtttaatgattttttgatagacatgctaatttaaaatatagatat
     ggaaatagaaagtttggtgtaaaggattttatgtagatacagtaggtagaaataaaaag
     gtaattgaaaattacattcgtaatcaattacaagaagatatcgttgcagatcaaatttca
     atggaagaatacctagatccctttacaggagaagaaattaaaaaaaagacgaaaaaaatag


50   Sequence 2932
     MRYAYGSEKVLPLSKKTSHHARIYIGQPTKIFNTRRCFFMSSDTNSLAHTKWNCKYHIVF
     VPKYRRQVIYGKIKRDIGVILRQLCERKGVEIIEAEASKDHIHMLVSIPPKLGVSSFVGY
     LKGKSSLMIFDRHANLKYRYGNRKFWCKGFYVDTVGRNKKVIENYIRNQLQEDIVADQIS
     MEEYLDPFTGEEIKKRRKK*
55
     Sequence 2933
     Contig_0517_pos_4627_4100
     is similar to (with p-value 1.0e-41)
      >gp:gp|AF006000|AF006000_5 Bordetella pertussis D-3-phospho

glycerate dehydrogenase homolog (serA) and Brgl (brgl) genes
, complete cds. NID: g2290988.
atggatcttactccaaatgaaatttataacttagttatatatcaattaggtgcgttaagt
ggcttttgtaaaatcaatcatgtaaaaatgatgcatgttaaacctcatggtgcccttttat
caaatggggggctagaaataaagaaattgcacatgcaattgctcaagcagtttttgatttc
gactcaaatctaattttcgtcggcttagcgaatacattacttatttcggaagctgaatta
gtggggcttaaggtagcttcggaagtatttgctgaccgtcgttatgaagatgacggacaa
ttggtaagtagaaaaaaaaccgatgccactatcactaatactgacgaagcaatccaacaa
gcattaaaaatggttttggaaaataaagttgtaagtaaaaatggaaaaatcatcgatttg
aaagctgatacaatttgtgttcacggagatggaaaacacgcattagaatttgttacgcaa
attagaaatgaattaatgaaagaaggcattgatattcaatccttatag


Sequence 2934
MDLTPNEIYNLVIYQLGALSGFCKINHVKMMHVKPHGALYQMGARNKEIAHAIAQꞬVFDF
DSNLIFVGLANTLLISEAELVGLKVASEVFADRRYEDDGQLVSRKKTDATITNTDEAIQQ
ALKMVLENKVVSKNGKIIDLKADTICVHGDGKHALEFVTQIRNELMKEGIDIQSL*


Sequence 2935
Contig_0517_pos_2645_1959
is similar to (with p-value 2.0e-52)
  >gp:gp|AF025380|AF025380_1 Salmonella typhimurium IS200 tra
nsposase (tnpA) gene, complete cds. NID: g2555163. >gp:gp|YO
99990|STFLGLIS2_2 S.typhi flgL gene, gene encoding putative I
S200 transposase and gene encoding putative RNAseE-like prot
ein. NID: g2765044. >gp:gp|Y09991|STIS2T157_1 S.typhi encodi
ng putative IS200 transposase, 1575bp. NID: g2765048. >gp:gp
|Z54217|STISFLIBC_3 S.typhimurium fli[B,C] genes and inserti
on sequence IS200. NID: g1150641. >gp:gp|U44749|STU44749_1 S
almonella typhimurium putative IS200 transposase gene, compl
ete cds. NID: g1177216. >gp:gp|AF093749|AF093749_2 Salmonell
a typhimurium strain LT2 NADP+-linked malic enzyme (maeB), p
artial cds; insertion element IS200 transposase, complete cd
s; eut operon, complete sequence; and unknown genes. NID: g3
885908. >gp:gp|L25848|STYIS200A_1 Salmonella typhimurium IS2
00 insertion sequence from SARA17, partial. NID: g439618.
atgataggaattattggagcaatggaagaagaagtgacgatttttaaagcgtaaattgaat
gatatgaatgaaataaatattgcgcatgttaaattttatgttggcaagctaaaccacaaa
gaggtggttttaacacaaagtggtataggtaaagttaatgcttctatctcaacgactttg
ttaatagaaaaatttaatccagaagtcgtcattaatactggatcagcaggtgcactagat
caaacactatctattggagatatattagtgagtaatcatgtattatatcatgatgctaat
gctacagcgtttggttatgaatatggacaaatacctcaaatgcctaaaacttatactact
gatcctactttgttgaaaaaaacaatgcatgtattagaacaacaacaactgaatggtaaa
gtaggtatgattgttagtggtgatagtttataggtagctcagaacagcgacaaaaaatt
aagcaacaatttccagaagctatggctgtcgaaatggaggcaactgcaattgcgcaaaca
tgttatcaatttaaagtaccatttatcgtaactagagctgtttctgatttagcaaacggt
aaagccgatatttctttttgaagaatttttagataaagcagctttatcatctagtgagaca
gtttcattattagtagaatcattataa


Sequence 2936
MIGIIGAMEEEVTILKRKLNDMNEINIAHVKFYVGKLNHKEVVLTQSGIGKVNASISTTL
LIEKFNPEVVINTGSAGALDQTLSIGDILVSNHVLYHDANATAFGYEYGQIPQMPKTYTT
DPTLLKKTMHVLEQQQLNGKVGMIVSGDSFIGSSEQRQKIKQQFPEAMAVEMEATATAQT
CYQFꞋVPꞘIVTRAVSDLANGKADISFEEFLDKAALSSSETVSLLVESL*


Sequence 2937
Contig_0518_pos_1926_964
is similar to (with p-value 1.0e-49)
  >sp:sp|P45510|DHAK_CITFR DIHYDROXYACETONE KINASE (EC 2.7.1.
29) (GLYCERONE KINASE). >gp:gp|U09771|CFU09771_2 Citrobacter

freundii DSM 30040 cyclopropane fatty acid synthase (cfa) g
ene, partial cds, dihydroxyacetone kinase (dhaK), glycerol d
ehydrogenase (dhaD), transcriptional activator (dhaR), 1,3-p
ropanediol dehydrogenase (dhaT), glycerol dehydratase (dhaB)
, glycerol dehydratase (dhaC) and glycerol dehydratase (dhaE
) genes, complete cds. NID: g1229153.
atgaaaaagttaattcaagataaaaacacaattttaaaagatatgcttgatggaattaca
gtttcaaacaacgatgttgaagttgtatctgacactattgttgttagaaagcataaaaaa
caatcaggtgttgcactcgtttctggggggcggcagtggacatgaacctgcacacgcagga
tttgtagcagaaggcatgctcgatgcagctgtatgtggagaaatcttcacttcacctaca
cctgataaaatattagatgccattaaagctgtggacaatggtgacggcgttctacttgtt
attaaaaactatgcaggagacgttatgaactttgaaatggctcaagaaatggctcaaatg
gaagatattaaagttgaaagtgttattgtcagagatgatattgctatttctgacccggaa
aaacgccgtggtgtagcgggtacagtatttgtgcataaatatgctggatacctagctgaa
aaaggtgttgcacttgatgaaatcaaatctaaagttgaggcacttttaccagatattaaa
agtattggtatggcattaacgcctccaatggtgcccacaactggtaaaaacggtttcgat
attgaagacaatcaaatggaaattggtatcggtattcacggtgaaaaaggtttacatcgt
gaagacgtacaacctattaacgtgattgttgaacgtttactcgatcaattatacaaagaa
attgagaaaaaaccctttaatcgtaatggttaatggtatgggtggtacgccactatcagaa
ttaaatatagttactaaatatctagatgaacaattcaatcagaatgatattggtgttaaa
caatggttcgtaggtgactatatgacagcgttagacatgcaaggcttctctataactgta
ctccccttcagtgaagaattgagtgaagctttagctgcacctacagcaagtaaatatttc
taa

Sequence 2938
MKKLIQDKNTILKDMLDGITVSNNDVEVVSDTIVVRKHKKQSGVALVSGGGSGHEPAHAG
FVAEGMLDAAVCGEIFTSPTPDKILDAIKAVDNGDGVLLVIKNYAGDVMNFEMAQEMAQM
EDIKVESVIVRDDIAISDPEKRRGVAGTVFVHKYAGYLAEKGVALDEIKSKVEALLPDIK
SIGMALTPPMVPTTGKNGFDIEDNQMEIGIGIHGEKGLHREDVQPINVIVERLLDQLYKE
IEKKPLIVMVNGMGGTPLSELNIVTKYLDEQFNQNDIGVKQWFVGDYMTALDMQGFSITV
LPFSEELSEALAAPTASKYF*

Sequence 2939
Contig_0519_pos_4309_3827
is similar to (with p-value 2.0e-25)
  >gp:gp|AF008183|AF008183_1 Populus balsamifera subsp. trich
ocarpa X Populus deltoides 4-coumarate:CoA ligase 2 (4CL2) m
RNA, complete cds. NID: g2911796.
atgaataaagtgctatttgagcaatatcctcattttgttaaaaatcacattcatattcct
gatggcttatcagaaaatctagaagcagaagcggaacgatataacaatttattagatgaa
agagggccaatcgatattcaaattttaggaattggagaaaatggtcacattggttttaat
gaaccagggactgacttcaatagtgaaacacatgtggtgaacttaacagaaagcaccata
aaagcaaatagtcgattttttgacaatgaaaaggatgttcctagacaagcagtttcaatg
ggggtaaaaagtattttaaaagcaaaaaggattatcctactcgcatttggtccaaagaaa
aaagaggctataagtaaactgttaaatgaacaggttaccgaagatgtacctgcgaccatt
ttacacacacaccctaatgttgaagtttatgtagacgatgaagcagcgccagattgttta
taa

Sequence 2940
MNKVLFEQYPHFVKNHIHIPDGLSENLEAEAERYNNLLDERGPIDIQILGIGENGHIGFN
EPGTDFNSETHVVNLTESTIKANSRFFDNEKDVPRQAVSMGVKSILKAKRIILLAFGPKK
KEAISKLLNEQVTEDVPATILHTHPNVEVYVDDEAAPDCL*

Sequence 2941
Contig_0519_pos_1586_288
is similar to (with p-value 2.0e-22)
  >sp:sp|Q04802|NAG1_CANAL GLUCOSAMINE-6-PHOSPHATE ISOMERASE
(EC 5.3.1.10) (GLUCOSAMINE-6- PHOSPHATE DEAMINASE). >pir:pir
|A46652|A46652 glucosamine-6-phosphate isomerase (EC 5.3.1.1

0) – yeast (Candida albicans) >gp:gp|L07558|YSANAG1A_1 Candi
da albicans glucosamine-6-phosphate deaminase (NAG1) mRNA, c
omplete cds. NID: g170885.
gtgtttgaagatcgacatctaacgtatggagaattaagtaaagaaatttatcaggctagt
atgcgctataaagaagtaaaattaaacaaaaaagtaggtctaatggatgaacatcctgta
aataatattattaactattttgcggtacatcaaagaggtggaattccttgcatttttaat
catcaatggagtaatgaaaggatacatcaacttgtaaaaagttatgacatacaatggtta
attaaagataatcatcttacctcaaatcatgataactcaatttataatgatgaggttatc
ccacgtaatgttatacatataggtttcacgtcaggaactacaggtttacccaaagcgttt
tatagaaatgaacattcttggatagtttcttttaaggaaaatgagaaattactccagcat
tgtgaagaaaccattgtagcaccgggtcctttatcacattcactttcattgtacgcatgt
atttatgcattaagtactggaaaaacatttataggtcaaaaaaattttaatccactatct
cttatgcgtcttattaatcaattgaacaaaacgacagcaatatttgtagtgccaacgatg
gtacaacaacttatttcaactcaacgacattgttcatcgattaaaagtattttgagtagt
ggtgctaaacttacattgcaacagtttcaacaaatcagaaatttatatccacaagcaaat
tt..at.ag.atttttttgggacatctgaagcaagtttt ataagctacaattttaacc..atca
tctcctgctaattctgttggtaaactttt ccctcatgtcgagacacgattattaa..tcaa
gatgatgatgcagtaggattattagccgttagaagtgaaatggtgtttagtggttatgtt
ggacaaagcaatcaagaggggcatggattaaaacaggcgacttcgcttatattaaaaat
caacatttgtttttagtaggtagagagagtgatcgtattatagttgggggattaatgta
tatccaacagctattgaaagcttaattatggatattgaaggcattgatgaggccttgtc
attggtataccacatgctaaatttggagaaatagcgatattgctttattcaggtaaagta
caattgaattaccgacaaattaaatctttttttaatgaaacatctttcaagacaagaagtt
ccatcaaaattaaagaaaattgaccatatgatttatacagaatcaggaaagattgctaga
aaagagatgaaaaataaatttattaatggagagttataa

Sequence 2942
VFEDRHLTYGELSKEIYQASMRYKEVKLNKKVGLMDEHPVNNIINYFAVHQRGGIPCIFN
HQWSNERIHQLVKSYDIQWLIKDNHLTSNHDNSIYNDEVIPRNVIHIGFTSGTTGLPKAF
YRNEHSWIVSFKENEKLLQHCEETIVAPGPLSHSLSLYACIYALSTGKTFIGQKNFNPLS
LMRLINQLNKTTAIFVVPTMVQQLISTQRHCSSIKSILSSGAKLTLQQFQQIRNLYPQAN
LIEFFGTSEASFISYNFNQSSPANSVGKLFPHVETRLLNQDDDAVGLLAVRSEMVFSGYV
GQSNQEGAWIKTGDFAYIKNQHLFLVGRESDRIIVGGINVYPTAIESLIMDIEGIDEALV
IGIPHAKFGEIAILLYSGKVQLNYRQIKSFLMKHLSRQEVPSKLKKIDHMIYTESGKIAR
KEMKNKFINGEL*

Sequence 2943
Contig_0523_pos_0_303
is similar to (with p-value 1.0e-31)
 >sp:sp|P14638|TRPB_METVO TRYPTOPHAN SYNTHASE BETA CHAIN (EC
 4.2.1.20). >gp:gp|M35130|MVOTRPBA_2 M.voltae tryptophan syn
thase operon (trp) genes, complete cds. NID: g150070.
atgaaaattcaaacagaagtagatgaattgggctttttcggtgaatatggtggccaatat
gtacctgaaacattgatgccagctattattgaacttaaaaaagcatatgaggacgcgaaa
tcagatactcacttcaagaaagaatttaattattatttaagtgaatatgttggtagagaa
acgcctttaacatttgctgaatcatacacaaaattgttaggtggtgccaaaatatatctt
aaaagagaagacttaaatcacactggtgctcataaaattaataacgcgataggacaggcA
TAA

Sequence 2944
MKIQTEVDELGFFGEYGGQYVPETLMPAIIELKKAYEDAKSDTHFKKEFNYYLSEYVGRE
TPLTFAESYTKLLGGAKIYLKREDLNHTGAHKINNAIGQA*

Sequence 2945
Contig_0527_pos_178_1065
 >sp:sp|P22983|PODK_CLOSY PYRUVATE,PHOSPHATE DIKINASE (?C 2.
7.9.1) (PYRUVATE,ORTHOPHOSPHATE DIKINASE).
atgggagcaacagcatttgtctataacggtcgtttccaccctgaaacatatctcgagtta
cttcaaaattatcaaattaatgttctatgttgtacaccaacagaatatcgtatgatggct

```
aaacttagtcatttagaacagtacaatttagagtatttacacagtgcggtgtctgcgggt
gaacctttaaatcgagaagttgttgaacaatttaaacgtcattttaatattactgttcga
gatggatatggacaaaccgaaagtacattgttgatcggatttctaaaagatactgaacca
cgtatgggttctatgggcaaaggtatacctggtagttttgttactgtcattgacgatgat
ggtaaagaggttggtccaaatgttaaaggtaatatcgccgtgcctttagacttaccggct
ttatttaaaggttactttaaagatgaagcacgcacaaaagcagcttcaacaggtgattat
tatgttactggagaccaagctcatattgataatgatggttatttctggttcgaaggtcgc
cgtgacgatattatcattagttcaggatataccattggacctttcgaggtagaagatgca
ctaacaaatcacgcagctgttaaagaatgtgcagttgttgcaagtcctcatgacattcgt
ggaaatattgttaaagcatttatcatcttgcaagatgattatgaagcaagtgatgagtta
atccaagaattacaagtattttgtaaaaatgaagtagcaccgtataaatatccaagagca
attgaatttgttgaacatctaccaaaaacaaattcaggtaagatacgtcgtgttgaatta
cgtgacgcagaaataaaaaaatataaacaacaagattcatcacattaa
```

```
MGATAFVYNGRFHPETYLELLQNYQINVLCCTPTEYRMMAKLSHLEQYNLEYLHSAVSAG
EPLNREVVEQFKRHFNITVRDGYGQTESTLLIGFLKDTEPRMGSMGKGIPGSFVTVIDDD
GKEVGPNVKGNIAVPLDLPALFKGYFKDEARTKAASTGDYYVTGDQAHIDNDGYFWFEGR
RDDIIISSGYTIGPFEVEDALTNHAAVKECAVVASPHDIRGNIVKAFIILQDDYEASDEL
IQELQVFCKNEVAPYKYPRAIEFVEHLPKTNSGKIRRVELRDAEIKKYKQQDSSH*
```

Sequence 2947
Contig_0527_pos_1455_0
is similar to (with p-value 3.0e-67)
 >gp:gp|AF068246|AF068246_1 Mus musculus SA protein mRNA, co
mplete cds. NID: g3928675.
```
atgaaagacttacttggtggtaaaggtgccaatctttcagagatgaagagactcggacta
ccagtaccagatggttttacaattacgactgaagcttgtattacatatttaaaacaaaat
gaagaactacctacagaagtaaagacacaattaattgatcatttagcagcttttctaaa
cgaacaggaaaagccttttcctctgatgataacttgttattagtatcagtacgtagtggt
gcaaaaatttccatgccaggaatgatggatacaattcttaatttaggacttaatgatgat
aatgtaaaaaagcttgttgacaaaacaaatgatgcacgatttgcatatgattgttaccgt
cgtttactacaaatgtttggagaggttgtttatggtattccaatgacagctttcgataca
tattttaatgattttaaaacaaagcatcgttatcaaaatgatgctgaaattccggcagaa
ggactacaaactatatgtgaaaaatataaagaaatctatgtagaagaggcatataaacct
tttccccaagaaccgttaaagcaattagaagaagcaattgaagcagtatttaaatcttgg
gataatgatcgtgcacgtgtatatagagatttaaatgatattccacatgatattggtaca
gccgtaaatattcaggaaatggtatttggtaatagtggtgaaaatagtggtacgggtgta
gcatttacaagaaatccagttactggagaaaatcatttattcggagagtacttacttaat
gctcaaggtgaagatgttgttgcaggcattcgtaccctaggatattgacactttaaaa
caacaaatgccagatgtacatcaagagtttgttgatgtaaccaaacaacttgaaaaacat
tacaaagatatgcaagacatagaatttacaattgaaatggtaaactttatttattacaa
acacgtaacggaaaacgtactgctaaagctgcaataaaaattgctgtggatttagttcac
gagcaattaattacacgtgaagaagcagtatcaaaagtagaggtaaaatcaatagaccaa
ttattacatcctaattttaatgaagaatcattaaagcaagcgacagtggtttctaaaatg
ggcttaccagctagtcccggtgcagcaacaggaaaagtagtcttctctgctgaagaagca
aaacttcaagctgaaaatggtaataaagtagtgttaatgcgaccggaaacatcaccatgaa
gatattgaggggatggtagcaagtgaagcaatcgtaacaactcatggtggtatgaatca
cacgctgctgttgtagcaagaggaatgggcaaatgttgtgtgacaggatgttcgaatgta
gagatagatacagtgaacaaaacagtatattatcctgaaggtgaattacatgaagggggat
atcgtttctgtagatggttcagctggtgatttatatttaggagcaattgaaacagtcaat
gctgaacatagtgaagagttcgatcaatttatgacttggtctgaagagattgcaagactg
caagttagaatgaatgctgaaacaccacaagatataaaagctggatataattttggttct
aaaggaataggtttagttcgtactgagcacatgttctttggccctgaacgtttaatagaa
atgcgccgttttatcttagcttcaaatcatgacgaacgtgtacaagctttagaaaaaatt
aaaacataccaagtagaagattttgaaacaattttcagattatctcaagatagacctacg
attgttcgtttacttgatccaccgttacatgagttcttaccatcatctgaagaagatata
aacaatgtttctcaacagctgaatgtatcttcagagttcttacgcaagcgaatcgttgac
ttaaatgaggtcaattcaatgcttggtcatcgtggttgtcgtttggctgtgacttatcca
```

```
gagttatatgagatgcaagttgaagctatcattgaaagtgttattaagcttcaaaaagag
ggcataacgtgcctaccagaaattatgattcctctcgtgtcaacagtagaagaatttaca
actttaaaagaacgattagttaatacaattacacatttagaaaaagaatcacaacaagat
atacaatatatgataggtactatgattgaaacgcctagagcatgcttgattgcgaatgac
cttgcgaaacattgtgatttcttcagttttggtactaatgatttaacgcaattgacattt
ggtttctctagagatgatgcaggaaaattcataaatgtgtatactgaaaataacatttta
cagcttgacccattccaaactttagatagagaaggtgtaggacgactaattcaattagct
gttgaacaagctaaaaatacaaatccagagataaaaattggtgtattta
```

10 Sequence 2948
MKDLLGGKGANLSEMKRLGLPVPDGFTITTEACITYLKQNEELPTEVKTQLIDHLAAFSK
RTGKAFSSDDNLLLVSVRSGAKISMPGMMDTILNLGLNDDNVKKLVDKTNDARFAYDCYR
RLLQMFGEVVYGIPMTAFDTYFNDFKTKHRYQNDAEIPAEGLQTICEKYKEIYVEEAYKP
FPQEPLKQLEEAIEAVFKSWDNDRARVYRDLNDIPHDIGTAVNIQEMVFGNSGENSGTGV
15 AFTRNPVTGENHLFGEYLLNAQGEDVVAGIRTPKDIDTLKQQMPDVHQEFVDVTKQLEKH
YKDMQDIEFTIENGKLYLLQTRNGKRTAKAAIKIAVDLVHEQLITREEAVSKVEVKSIDQ
LLHPNFNEESLKQATVVSKMGLPASPGAATGKVVFSAEEAKLQAENGNKVVLMRPETSPE
DIEGMVASEAIVTTHGGMTSHAAVVARGMGKCCVTGCSNVEIDTVNKTVYYPEGELHEGD
IVSVDGSAGDLYLGAIETVNAEHSEEFDQFMTWSEEIARLQVRMNAETPQDIKAGYNFGS
20 KGIGLVRTEHMFFGPERLIEMRRFILASNHDERVQALEKIKTYQVEDFETIFRLSIDRPT
IVRLLDPPLHEFLPSSEEDINNVSQQLNVSSEFLRKRIVDLNEVNSMLGHRGCRLAVTYP
ELYEMQVEAIIESVIKLQKEGITCLPEIMIPLVSTVEEFTTLKERLVNTITHLEKESQQD
IQYMIGTMIETPRACLIANDLAKHCDFFSFGTNDLTQLTFGFSRDDAGKFINVYTENNIL
QLDPFQTLDREGVGRLIQLAVEQAKNTNPEIKIGVFX

25

Sequence 2949
Contig_0528_pos_3612_3271
 No hits found
atgtcaaatagcgcttcaatatctcagagcaacgtcgcaagtcaaagtactacagcaagt
30 ttgagccaatctgaatcagcaaatgattcaatgagttcatctctgtccgagtctaactca
ataacatccgaaagtaatacaaatagcaaatcggaaattgaatcaaaaagtacgtctaca
agcgagttcttgtcagaatcaggaagtgtatctaactcagaaaaatctgagtcaatttct
cattctcaatcaacatcagctacaccttcttctcaatcgacttaccaacaacaacctaaa
gaagagaagaaaggtttctttgcacgtctatttaacttataa
35

Sequence 2950
MSNSASISQSNVASQSTTASLSQSESANDSMSSSLSESNSITSESNTNSKSEIESKSTST
SEFLSESGSVSNSEKSESISHSQSTSATPSSQSTYQQQPKEEKKGFFARLFNL*

40 Sequence 2951
Contig_0530_pos_2171_1191
is similar to (with p-value 3.0e-28)
 >gp:gp|Y14325|ATY14325_1 Arabidopsis thaliana mRNA for meva
lonate diphosphate decarboxylase. NID: g2288886. >gp:gp|AC00
45 5499|ATAC005499_10 Arabidopsis thaliana chromosome II BAC T6
A23 genomic sequence, complete sequence. NID: g3785992. >gp:
gp|Y17593|ATH17593_1 Arabidopsis thaliana MVD1 gene, exons 1
to 9. NID: g3250735.
gtgaaaagtggcaaagcacgagcacatacaaatattgcgttgattaagtattgggggaaa
50 gctgatgaaacttacattattcctatgaataatagtttatcagttaccttagatagattt
tatactgaaacaaaagtgacatttgaccctgatttttactgaagattgccttattttaaat
ggtaatgaagtgaatgccaaagagaaagaaaagattcaaaactatatgaatatagtgaga
gatttggctggaaatcgtttgcatgcgcgaattgaaagtgaaaattatgtgccaacagca
gcaggacttgcttcttcagcgagtgcttacgctgctttagctgccgcttgtaatgaagct
55 ttgtcattgaacttatcagatacagacttatcacgattagctcgacgtggttcaggttct
gcttctagaagtattttttggtggatttgccgaatgggaaaaagggcatgatgatttaact
tcatatgcacatggtattaattccaatggttgggaaaaagatttatcaatgatatttgta
gtcattaacaatcagtcaaaaaaagtatctagtaggtcaggaatgtcactaacaagagat
acttctagattttatcaatattggttggatcacgttgatgaagatttaaatgaagaaaa
```

```
gaggcagtcaaaaatcaagattttcaacgcttaggagaagtcattgaagcaaatggttta
cgtatgcatgccactaacttaggcgctcaacctcctttcacgtatttagtgcaagaaagc
tacgatgctatggcgattgtggaacagtgtcgaaaagccaatttaccttgttactttaca
atggacgcgggtcccaatgtaaaagtttttagtagaaaagaaaaataaacaagctgtgatg
gaacaattttttaaaagtatttgacgaatcgaagattatagcaagtgatatcattagctct
ggtgttgaaattattaagtaa
```

Sequence 2952
VKSGKARAHTNIALIKYWGKADETYIIPMNNSLSVTLDRFYTETKVTFDPDFTEDCLILN
GNEVNAKEKEKIQNYMNIVRDLAGNRLHARIESENYVPTAAGLASSASAYAALAAACNEA
LSLNLSDTDLSRLARRGSGSASRSIFGGFAEWEKGHDDLTSYAHGINSNGWEKDLSMIFV
VINNQSKKVSSRSGMSLTRDTSRFYQYWLDHVDEDLNEAKEAVKNQDFQRLGEVIEANGL
RMHATNLGAQPPFTYLVQESYDAMAIVEQCRKANLPCYFTMDAGPNVKVLVEKKNKQAVM
EQFLKVFDESKIIASDIISSGVEIIK*

Sequence 2953
Contig_0532_pos_1713_3254
 >sp:sp|P54715|PTIB_BACSU PTS SYSTEM, ARBUTIN-LIKE IIBC COMP
ONENT (PHOSPHOTRANSFERASE ENZYME II, BC COMPONENT) (EC 2.7.1
.69). >gp:gp|Z99108|BSUB0005_89 Bacillus subtilis complete g
enome (section 5 of 21): from 802821 to 1011250. NID: g26330
55. >gp:gp|D50543|D50543_3 Bacillus subtilis DNA for 76-degr
ee region, complete cds. NID: g1486240.

```
atgtttgctttttttcggtattgttttgggattcgctacattatttaaaaatccaaccatt
atgggaggattagctgatcagcaaacattttggtttaaattttggtctgttattgaatca
ggtggttgggtaatatttacacatatggaaattgtctttgtagttggcttaccattatct
cttgctaaaaaagcaccaggacatgcagctttagcagctctaatgggatatttaatgttt
aatacttttatcaatgcaattttaactcaatggccacatacttttggcgctaatttaaaa
aaaggtgtagaaaacacaacaggattaaaatcgattgcaggtattgaaacgttagatacc
aatattttaggtgcaatcattatctcaggaataataacgtggatacataatagatattac
agtaagcgtttacctgaaatgttaggtgtatttcaaggattaacattcgttgtaacaatc
tctttctttgtcatgctaccagtggcagcaatcacttgtgttgtttggccaacgattcaa
cacggtattgcttcaattcaatattttattgttgcatcaggttatataggtgtttggtta
tatcatttcttagagcgtgtactaatacctacaggattgcatcattttatctatgcacca
atcgaagtaggtccagtagttgttaatcatggtttgaaagcagaatggcttcaacactta
aaccagtttgccgaaagtaataaaccacttaaagaacaattcccatatggatttatgttg
caaggaaatggaaaggttttttggttgtctaggtatagcattggcaatgtatgcaactaca
ccaaaagaaaatcgtaaaaaagttgctgcattattaataccagcaacacttacggcagta
gtagctggtattacagaaccacttgaatttacattcttatttattgcgccatttttattc
gtattacatgcactactagcagcaactatggatacactgatgtatggatttggtgttgta
ggtaatatgggtggcggtgtactagatttttattgcaactaactggataccattaggaaaa
gcacattggatgacatatgtctttcaagtagtaattggtttaatctttgttgcaatttac
tacttcttatttaaatatttaattttaaaatttgatattccattaccaggacgtaagaaa
ggtgaagaagaagttaaattattttccaaacaagattataaagataaaaaaggagattca
actcgcaatcattcacctaatagtgaatatgaagaaaaggcgatgtactacctagaaggt
cttggaggaaaagaaaatattaaagacgttacaaattgtacgacacgtctacgtttaact
gttaaagacgaaagtaaagttcaagaaagtgcctattttacacataatcaaatgtctcat
ggtttagtaaaaagtggcaaaagtgttcaagtcgtcgttggaatgtctgtacctcaggta
agagaagcatttgaaaatattgtaaatgatgatctatcttag
```

Sequence 2954
MFAFFGIVLGFATLFKNPTIMGGLADQQTFWFKFWSVIESGGWVIFTHMEIVFVVGLPLS
LAKKAPGHAALAALMGYLMFNTFINAILTQWPHTFGANLKKGVENTTGLKSIAGIETLDT
NILGAIIISGIITWIHNRYYSKRLPEMLGVFQGLTFVVTISFFVMLPVAAITCVVWPTIQ
HGIASIQYFIVASGYIGVWLYHFLERVLIPTGLHHFIYAPIEVGPVVVNHGLKAEWLQHL
NQFAESNKPLKEQFPYGFMLQGNGKVFGCLGIALAMYATTPKENRKKVAALLIPATLTAV
VAGITEPLEFTFLFIAPFLFVLHALLAATMDTLMYGFGVVGNMGGGVLDFIATNWIPLGK
AHWMTYVFQVVIGLIFVAIYYFLFKYLILKFDIPLPGRKKGEEEVKLFSKQDYKDKKGDS
TRNHSPNSEYEEKAMYYLEGLGGKENIKDVTNCTTRLRLTVKDESKVQESAYFTHNQMSH
```

GLVKSGKSVQVVVGMSVPQVREAFENIVNDDLS*

Sequence 2955
Contig_0533_pos_2932_4371
is similar to (with p-value 7.0e-20)
 >gp:gp|AB005556|AB005556_1 Plectonema boryanum DNA for NADP
H:protochlorophyllide oxidoreductase, hypothetical protein,
partial and complete cds. NID: g3123724.
atggaggtgtttaacatgacaaatcaattatttattaacaatgaatttatagaaagtcag
tctaaagagacaatggatgtcattaatccagctactggcgaggcatttgatactatcact
cttgcaactgaagaggaagtaaacgacgccattgaaaaatcgcaacaagcacaacttgaa
tgggagcgcgtgcctcaacctacacgtgcggaacatgttaaattacttataccttatta
gaaaaaaatcgcgatgaaatagctcaattatacgtaaaagaacaaggtaaaactttagcg
caagcttatggagaaattgacaaatcaatctcatttatcgattatatgacaagtctgagt
atgtcagataaaggacgtgttctacaaaatagtattgcaaatgaaacgattcaaattatc
aacaaacctatcggagttactgctggtattgtgccatggaacgcaccgatacttgtcctt
atgcgaaaagtcattccagctatagtaactggttgttcagtagtgattaaacctagtgaa
gagacaacgttactcactcttcgattagctgaattattcagagcatcaactataccagca
ggattgtttcaaattgttcctggcactggagaaacagtaggtacacaattagcttcgcat
aaagacattcaacttatttctttaactggaagtatgagagctggtaaatctgtttacgaa
aatgctgctcaaactgtaaagaaagtaaatttagaattaggtggaaacgcaccagtcatt
gtcacatcaaatgccgatttagataaagcagttgactatatcgtgacagcccgtataaat
aatgcaggtcaagtttgtacgtgccctgaacgtatctttgtacatgaagatgttcacgat
gactttttaaataaagtaacttccaaaatgaaaagcttaactgttggagatccatttgat
gaaaacaccgattacggcgcaattattaaccaaaaacaacttgatagtattcatgaaaag
gttcaagatgctattaaaaatggtgcaacattgatgactggtggacatcaattaaaacgc
catggtttcttctacgcaccaacagtattagataacgttagaaaagacgataatgtgttt
aaagatgaaatttttggtccagttcttgcgataacaacctaccgtgatattgaacaagtg
attgaagacgctaatgatacaaacgctggcttatcttcttatatcttctctgaaaattta
acagaagtaatgacagcaaccgaacgtctaaaatttggtgaagtatatgcaaattgtgag
gctgaagaagtggttaatggctatcacgcaggttggcgtgaatcaggcttaggtggcgct
gacggtattcacggttttgaagaatactacaataccacagtaagttatattagatactaa

Sequence 2956
MEVFNMTNQLFINNEFIESQSKETMDVINPATGEAFDTITLATEEEVNDAIEKSQQAQLE
WERVPQPTRAEHVKLLIPLLEKNRDEIAQLYVKEQGKTLAQAYGEIDKSISFIDYMTSLS
MSDKGRVLQNSIANETIQIINKPIGVTAGIVPWNAPILVLMRKVIPAIVTGCSVVIKPSE
ETTLLTLRLAELFRASTIPAGLFQIVPGTGETVGTQLASHKDIQLISLTGSMRAGKSVYE
NAAQTVKKVNLELGGNAPVIVTSNADLDKAVDYIVTARINNAGQVCTCPERIFVHEDVHD
DFLNKVTSKMKSLTVGDPFDENTDYGAIINQKQLDSIHEKVQDAIKNGATLMTGGHQLKR
HGFFYAPTVLDNVRKDDNVFKDEIFGPVLAITTYRDIEQVIEDANDTNAGLSSYIFSENL
TEVMTATERLKFGEVYANCEAEEVVNGYHAGWRESGLGGADGIHGFEEYYNTTVSYIRY*

Sequence 2957
Contig_0533_pos_8501_7722
is similar to (with p-value 2.0e-98)
 >sp:sp|P25553|ALDA_ECOLI ALDEHYDE DEHYDROGENASE A (EC 1.2.1
.22) (LACTALDEHYDE DEHYDROGENASE).
gtgggttacggtgtagcacaagcaggtgctgaacgtatcgttcctacaacattagaacta
ggtggtaaaagtgctaatattatctttgatgatgctaatttagagcaagtgattgaaggt
gttcaacttggtatattatttaaccaaggtgaagtatgtagtgcaggttcaagattatta
gtacaatcatctatttacgatgaattgttgccaaaattgaaagaagcatttgaaaatatt
aaagttggcgatccatttgatgaagatactaaaatgagtgcgcaaactggaccagaacaa
ttagataaaattgaaagttatataaaaattgctgaagaagatgacaaagcgaacatctta
actggtggtcatcgaatcacagataacggcttagacaaaggttacttctttgagcctaca
attattgagattaacgataacaagcatcaacttgctcaagaagaaatcttcggtccagtt
gtagtagttgaaaaattcgatgatgagcaagaagctatcgaaattgcaaatgattctgag
tatggtttagctggaggtatcttcactacagatattcatcgtgcattaaatgtagctaaa

gctatgagaacaggtcgtatttggattaatacttataatcaaattcctgctggtgcgcca
ttcggaggatataaaaaatcaggtattgggcgcgaagtatataaagatgctatcaaaaac
tatcaacaagttaaaaatatctttattgatacaagcaaccaaactaaaggtttatattaa

5   Sequence 2958
VGYGVAQAGAERIVPTTLELGGKSANIIFDDANLEQVIEGVQLGILFNQGEVCSAGSRLL
VQSSIYDELLPKLKEAFENIKVGDPFDEDTKMSAQTGPEQLDKIESYIKIAEEDDKANIL
TGGHRITDNGLDKGYFFEPTIIEINDNKHQLAQEEIFGPVVVVEKFDDEQEAIEIANDSE
YGLAGGIFTTDIHRALNVAKAMRTGRIWINTYNQIPAGAPFGGYKKSGIGREVYKDAIKN
10  YQQVKNIFIDTSNQTKGLY*

Sequence 2959
Contig_0533_pos_0_1310
is similar to (with p-value 7.0e-62)
15  >sp:sp|P40047|DHA3_YEAST ALDEHYDE DEHYDROGENASE, MITOCHONDR
IAL 3 PRECURSOR (EC 1.2.1.3). >gp:gp|U56605|SCU56605_1 Sacch
aromyces cerevisiae precursor aldehyde dehydrogenase gene, n
uclear gene encoding mitochondrial protein, complete cds. NI
D: g1336077.
20  gtgtttaatatgtcattggaattgcccgtgatgctaacgattgtgttgttcttagcacta
ggcattttagtcaatggttagcgagtagaataaaatggccatcgattgttgtcatggcc
atcgtaggtttacttgtaggacctattttttggattagcaaatccaaaagaggcacttgga
cctgaggcatttagttcaattgtatctcttgctgtagcaattatattatttgaaggtagt
agtaatctagatttagagaattaaaaaggcatttctaaagctgttataagaattattaca
25  ataggagcgggaattgcatggatttttaggagcaatcgctttacatgtcactatgaatttc
cctctgtctatttcatttgttatcggagggctattcctaatcactggtccaaccgtgatt
caacccttgctgaagcaagcgaaagtaaaaagaaatgtcgattcagtattaagatgggaa
agtattatcttagatcctattggacctattattgcactcactgcattttatgtattccag
atttttgaggaagggataggctttgtcgttattattctctttatcttgaagctcttagct
30  gcaattttaataggtttttggtgcagcctttctatttaattggcttataggtcaagataaa
attcctcaaagtttaatgccacccattcagttgtatttattcttttaacatttagtatc
tgcgatgagatattatcagaatcaggttactagctgtaacaatttttgggttaatgatg
gcgcgtaaaaaacgtcacgatctcatctttaaagaatcagaccactttatagataatgca
tcatccatacttgtgagtactgtatttattttaattacgtcttctctcactaaagatgtg
35  ttgctcaatgtgttatcttggcagctcatactctttagttggttatgattgtattagta
agaccaatttcagtgtttctttcaacattaggtactgaaataactaaaaaggaacgtgca
gtagtagcactaatggcgcctagaggtattgtggttttaacagtggcacaattcttctca
agtttatttatggacgataaaattcctatggctcaatatattacaccagtaactttggt
cttgtatttatcactgtagtcatttatggatttggtttacacctttaagtaaactgttc
40  ggtgtagcaagtacggagccaccaggcgtaatcattgtcggagaaagtgaatttttcgttc
catcttggtattaatctaagggatcatggtatacccgtcatgatgttcaa

Sequence 2960
VFNMSLELPVMLTIVLFLALGIFSQWLASRIKWPSIVVMAIVGLLVGPIFGLANPKEALG
45  PEAFSSIVSLAVAIILFEGSSNLDFRELKGISKAVIRIITIGAGIAWILGAIALHVTMNF
PLSISFVIGGLFLITGPTVIQPLLKQAKVKRNVDSVLRWESIILDPIGPIIALTAFYVFQ
IFEEGIGFVVIILFILKLLAAILIGFGAAFLFNWLIGQDKIPQSLMPPIQFVFILLTFSI
CDEILSESGLLAVTIFGLMMARKKRHDLIFKESDHFIDNASSILVSTVFILITSSLTKDV
LLNVLSWQLILFSLVMIVLVRPISVFLSTLGTEITKKERAVVALMAPRGIVVLTVAQFFS
50  SLFMDDKIPMAQYITPVTFGLVFITVVIYGFGFTPLSKLFGVASTEPPGVIIVGESEFSF
HLGINLRDHGIPVMMFX

Sequence 2961
Contig_0534_pos_7490_6273
55  is similar to (with p-value 3.0e-99)
>sp:sp|P39312|CYCA_ECOLI D-SERINE/D-ALANINE/GLYCINE TRANSPO
RTER. >pir:pir|S56433|S56433 hypothetical protein o470 - Esc
herichia coli >gp:gp|U14003|ECOUW93_120 Escherichia coli K-1
2 chromosomal region from 92.8 to 00.1 minutes. NID: g126317

2. >gp:gp|AE000492|AE000492_4 Escherichia coli K-12 MG1655 s
ection 382 of 400 of the complete genome. NID: g1790649.
atgttcatgagagctatgggagaattactgttatccaatttaggatttaaatcgtttggt
gacattgctcatcatcatattggttctatggcaggttttatggtgggggtggacatattgg

5    ttaacatggattatttcaggaatggcagaagtgactgctgttgccaagtatgtttccttc
tggtatccaacaattccaaactggttaacagctgcagcgactattttagtttttagttgct
ttaaatctattcagtgctaaattatttggagaattagaattttggctatctattattaaa
gttttgactattttagctttgatagccgttggtgttgttatgattgtatttggaatgaag
acaagctatggccctgcaacggtaacgaatatatggaaagacggaggcttttttccctaat

10   ggtgcacaaggtttcttcatgtcattccaaatggcaatttttctatttattggtattgag
ttgattggaataactgcaggggagactaaagatcctcacaaaacaattcctcaagcaatt
aataatgtaccgtttagaatattattattttatataggatcgttggcagttatcatgtct
gttgtaccatggcaacaattgaatcctgctgacagtccatacgttaaaatgtttggatta
gttggaatcccttttgcagcaggtattattaactttgttgtacttacagctgcagcctct

15   tcttgtaatagtggtatatttgctaatagccgtacgatgtttggattagctggaagaaag
caaggtccagcattcttacatagaaccaataagcacggcgtaccacattatgctatttta
gtgacatgtggcttattaagtatttcagtcgtgttaaatgcaatttttaaagatgcgact
aaagtgttcgtacaaattacaacattttcaactgtttttaaatattatgatttggacaatt
attatgatcgcgtatctaggttatttaagacatgaaccgaaacagcataaagaaagtaac

20   tataaaatgtggggcggaaaatacatggcttacagtattttaggggttctttgcatttatt
tttattatactattgattaatagtgcaacgcgttatgccgtactttctgcacccgtatgg
tttgttatcatgctattgatgtatcaaaaatataaaaaagaatctcgcaaagctaaaatt
aaaaatgaggaagagtaa

25   Sequence 2962
MFMRAMGELLLSNLGFKSFGDIAHHHIGSMAGFMVGWTYWLTWIISGMAEVTAVAKYVSF
WYPTIPNWLTAAATILVLVALNLFSAKLFGELEFWLSIIKVLTILALIAVGVVMIVFGMK
TSYGPATVTNIWKDGGFFPNGAQGFFMSFQMAIFSFIGIELIGITAGETKDPHKTIPQAI
NNVPFRILLFYIGSLAVIMSVVPWQQLNPADSPYVKMFGLVGIPFAAGIINFVVLTAAAS

30   SCNSGIFANSRTMFGLAGRKQGPAFLHRTNKHGVPHYAILVTCGLLSISVVLNAIFKDAT
KVFVQITTFSTVLNIMIWTIIMIAYLGYLRHEPKQHKESNYKMWGGKYMAYSILGFFAFI
FIILLINSATRYAVLSAPVWFVIMLLMYQKYKKESRKAKIKNEEE*

Sequence 2963
35   Contig_0539_pos_0_15071
is similar to (with p-value 8.0e-35)
 >gp:gp|X81475|MHLMP_1 M.hominis lmp1 and lmp2 genes. NID: g
587470.
atgggcacgttaaaatcattagttgctaaacaacctacagtacaaaaaacaagtgtttat

40   attaacgaagatcaacctgagcaatctgcctacaatgattccattacaatgggacaaact
ataattaataaaacagctgatccagtacttgataaaactttagttgataacgcaatcagt
aacatttcaactaaagagaatgcactgcatggtgaacaaaaattaacaactgctaaaacg
gaagcaattaatgcacttaatacattagctgatttaaacacacctcagaaagaggctatt
aaaacagctattaacactgctcatacaagaactgatgtaactgcagagcaaagtaaggct

45   aatcaaataaatagtgcaatgcacacgttgagacaaaacatttctgacaacgaatcagta
acaaacgaaagtaattatattaacgctgaacccgaaaaacaacatgcctttactgaggct
ctaaataatgctaaagaaatagttaatgaacaacaagccactcttgatgccaattcaatt
aaccaaaaagcacaagcgattcttactactaaaaatgctttagatggtgaagaacaatta
cgtcgtgctaaagaaatgccgatcaagaaatcaatacgttaaatcaattgactgatgcg

50   caaagaaatagtgaaaaaggtttagtcaacagttctcaaactagaacagaagttgcttct
caattagcaaaagctaaagaactaaataaggtgatggaacaactgaataaccttatcaat
ggtaaaaaccaaatgataaatagcagtaaatttatcaatgaagatgcgaaccaacaacaa
gcatattcaaatgcgattgcaagtgcagaagtgcttaaaaacaaatcacaaaaccctgaa
ttagataaagtaacaattgaacaagcaattaataatattaattctgcaattaataatcta

55   aacggtgaagctaaactgactaaagctaaagaagatgctgttgcttcaataaacaaccta
agcggattaactaacgagcaaaaaacaaaagaaatcaagccgttaatggctctcaaact
agagaccaagttgctaacgttttacgtgattcaaaggcattagatcaatctatgcaaaca
ttacgtgatttagttaacaatcaaaatgtaatacattcaacaagtaattattttaacgag
gattcaactcaaaagaatacttatgataatgcaattgataatggctcgacatatataact

```
      ggtcaacacaattcagaattaaataaatctactattgatcaaacgattagccaaattaac
      acagctaaaaatgatttacatggtgcagaaaagttacaaagagataagggaactgctaat
      caagaaattggacaattaggttatttaaatgaccctcaaaaatctgctgaggaatcctta
      gtcaacggttcaaatacacgttctgaagtagaagagcatcttaatgaagctaaatcatta
  5   aataatgcaatgaaacaattaagagataaagtagctgaaaagactaatgtcaaacaaagt
      agcgattacattaatgattcaactgaacatcaacgtgggtatgatcaagcacttcaagaa
      gcagaaaatattattaatgaaatcggtaatccaacattaaataaatcggaaattgaacaa
      aagttacaacaattgactgacgctcaaaatgcgttacaaggttcacatctattagaagaa
      gctaaaaataatgcgattactgaaatcaataaacttacagcattaaatgatgcacaacgt
  10  caaaaagcaattgaaaatgttcaagcacagcagacaatcccagcagttaatcaacaatta
      actttggatagagaaataaatactgcaatgcaagctttacgagataaagtaggccaacaa
      aataacgttcaccaacaaagtaattatttcaatgaagatgaacaaccaaaacataactat
      gataattctgtacaagccggtcaaactattattgataaacttcaagatccaatcatgaac
      aaaaatgaaattgagcaggctattaatcaaatcaatacgactcaaacagcgttaagtgga
  15  gaaaataaattacacactgaccaagaaagcacaaatagacaaatagaaggtttatctagt
      ttgaacacagctcaaatcaacgccgaaaaagatttagtcaatcaagctaaaacaagaaca
      gatgttgctcaaaagttagctacagctaaagaaataaattctgctatgagtaatttaaga
      gatggcattcaaaataaagaggacatcaaacgtagcagtgcatatatcaacgcagatccg
      actaaagttacagcttacgatcaagcactacagaacgcagaaaatatcatcaatgccaca
  20  ccaaacgtagagcttaataaagctacaattgaacaagcgctatcacgcgttcaacaagca
      caacaagatcttgatggtgttcaacaattagctaatgctaaacaacaagctacacaaact
      gtcaatgggttaaatagcttaaatgacggtcaaaagcgtgaattaaatctattaattaat
      tcagctaatacccgtacaaaagtacaagaagaattaaacaaagcaactgaatcgaaccat
      gcgatggaagctttaagaaacagtgttcaaaacgttgatcaagtaaaacaaagtaacaat
  25  tatgtcaatgaagatcaacctgaacagcacaattatgataatgctgtcaatgaagctcaa
      gctacaatcaacaacaatgctcaacctgttctgacaaattagctatagaacgtttaact
      caaactgttaacactacaaaagatgcattacatggtactcaaaaactgatacaagaccaa
      caagctgctgaaactggaatacgtggtttaacgagtctcaatgaacctcagaaaaatgct
      gaagtagctaaagtaactgcagcaacaacacgtgatgaagtgagaaatattcgtcaagaa
  30  gcaacaacattagatactgcaatgcttggtttacgtaaaagcattaaagataaaaacgat
      actaaaaatagtagtaaatatattaatgaggatcatgaccaacaacaagcttatgacaat
      gctgtaaataatgctcaacacgttatcgatgaaactcaagcaacgttaagctcagataca
      atcaatcaattggcaaatgccgtaactcaagctaaatctaatcttcatggagatactaaa
      ctacaacacgataaagatagtgctaaacaaacgattgctcaattacagaatttgaattca
  35  gctcaaaaacatatggaagattctttaattgataatgaatctacacgtacgcaagtccaa
      cacgatttaacagaagctcaagctttagatggtttaatgggtgccttaaaagaaagtatt
      aaagataatactaatattgtttcaaacggtaattacatcaatgcggaaccatctaagaaa
      caagcatatgatgcagctgtacaaaatgctcaaaatataataaatggaacgaatcaacca
      acaattaataaaggtaatgtcactacagcaacacaaaccgtgaaaaatactaaagatgcc
  40  ttagacggtgatcatagattagaggaagctaaaaataatgccaatcaaacaatcagaaat
      ctatctaatttgaacaatgcccaaaaagatgcagagaaaaatctagttaatagcgcatca
      acattagaacaagttcaacaaaacttacaaaccgctcaacaattagataatgctatgggt
      gagttacgacaaagtattgctaacaaagatcaagtgaaagcagatagtaaatatctaaat
      gaagatcctcaaattaagcaaaactatgatgatgcagtcaacgtgttgaaactattatt
  45  aacgaaactcaaaaccctgaattacttaaagcaaacattgaccaagcaactcaatccgtt
      caaaatgcagaacaagctttacatggtgctgaaaaattaaatcaagacaaacaaacgtct
      tcgacagaactagatggattaacagatttaacagatgcacaacgtgaaaaactcagagaa
      caaattaacacttctaatagtagagatgatattaagcaaaaattgagcaagcaaaagca
      ctaaatgacgcaatgaaaaaacttaaagaacaagttgcgcaaaaagatggtgttcatgct
  50  aacagtgattatacaaatgaagattctgcacaaaaagatgcgtataataatgcacttaaa
      caagcggaagacattattaataacagctcaaatcctaacttaaatgcacaagacattact
      aatgctttaaataatattaaacaagcacaagataaccttcatggagctcaaaaattacag
      caagacaaaaatacaactaatcaagccattggtaacttaaatcatcttaatcaacctcaa
      aaagatgcgcttatacaagctattaatggagctacatctagggaccaagttgcagaaaaa
  55  cttaaagaggccgaagcgcttgatgaagctatgaaacaacttgaagatcaagtgaatcaa
      gatgatcctaatttcaaatagcagcccattcataaatgaagactcagacaaacaaaaact
      tataatgataaaatccaagctgcaaaagaaataattaatcaaacatctaatccaacctta
      gataaacaaaaaattgctgatacacttcaaaatattaaagatgcagtgaataatttacat
      ggtgatcaaaaattagctcaatctaaacaagacgctaataatcaattaaatcatttagat
```

```
     gacttaaccgaagaacaaaaaaaccattttaaaccgttaattaataatgctgatactcga
     gatgaggtaaatàaacaactagagattgctaaacaattaaatggtgatatgagtacactt
     cataaagtcataaatgataaagatcaaattcaacatttaagcaattacattaatgctgat
     aatgataaaaaacaaaattatgataatgctattaaagaagctgaggatttaattcataat
 5   catccagatacattagatcataaagcattacaagatttattaaacaagatagaccaagcg
     cataacgaattaaatggagaatccagatttaaacaggctttagacaatgcttttaaacgac
     atagatagcttaaacagtctcaatgttccacaacgtcaaactgttaaggataacatcaac
     catgtgacaactctagaaagtttagctcaagaattgcagaaagcaaaagagcttaaʰgat
     gctaˈtgɛaagcaatgagagatagcattatgaatcaagagcaaattcgtaaaaatagcaat
 10  tatactaatgaagacttagctcaacaaaatgcctataatcatgcagtagataatataaat
     aacattattggtgaagacaatgcgacgatggatcctcaaataatcaaacaagcaactcaa
     gatataaatacagctataaatggattaaatggagatcaaaaacttcaagatgcaaagaca
     gatgctaaacaacaaattactaactttactggtttaactgaaccacaaaaacaagcattg
     gaaaacatcattaaccaacaaacaagtagagcaaatgttgctaaacagttaagtcatgct
 15  aaattcttaaatggaaaaatggaagaattaaaagttgcagtagccaaagcgtcattagta
     agacaaaatagtaactatattaatgaagatgtctctgaaaaagaagcatatgaacaagct
     attgcaaaaggtcaggaaataattaattcagaaaataatccaacaataagtagtactgat
     atcaatcgtaccattcaagaaattaatgatgctgaacaaaatcttcatggtgaaaataaa
     ttaagacaagcacaggaaattgcaaagaatgaaatacaaaatctagacggattaaaattca
 20  gctcaaataacaaaaattaatccaagatataggcagaacaacaactaaacctgcagtaact
     cagaaactagaagaagcaaaagcaataaaccaagctatgcaacaacttaaacaaagtata
     gccgataaggatgctactctaaattctagtaactatctcaatgaagattctgagaaaaag
     ttagcgtacgataatgctgtaagccaagctgaacaactcataaatcaacttaacgaccca
     actatggatataagtaatattcaagctattactcaaaaggtcattcaagcaaaagattca
 25  ttgcacggtgcgaataaacttgcacaaaatcaagcagattcaaatttaataataaatcaa
     tcaacaaatttaaatgataaacaaaagcaagcattaaatgacttaattaatcatgctcaa
     actaaacagcaagtggcagaaataattgcacaagctaataagttaaataacgaaatʒggc
     acɛɔʰɜəʰaacactcgtagaagaacagtcaaacgttcatcaacaaagtaaatataˈtaat
     gaagatɔcgcaagttcaaaatatttataatgactccattcaaaaaggtcgagaaatatta
 30  aacggcactacagatgatgtttaaacaacaataaaatagcagatgccattcaaaacatt
     catttaactaaaaacgatttacatggtgatcaaaaattacaaaaagcacaacaagatgca
     accaatgaattaaactattttaacaaatctaaacaattctcaaagacaaagcgagcatgat
     gagattaactctgctccttcaagaactgaagtttctaatgatttaaatcatgctaaagca
     cttaatgaagctatgcgtcaacttgagaatgaagttgctcttgaaaacagtgttaaaaaa
 35  ttaagcgactttatcaatgaagatgaagcggcacaaaatgaatatagtaatgcacttcaa
     aaagctaaagacattatcaacggcgttccaagtagcactttagataaagctacaattgaa
     gatgctttattagaattgcaaaatgctagagaaagtttacatggtgagcaaaaacttcaa
     gaggctaaaaatcaagctattgctgaaattgataatttacaagcattaaatcctggacag
     gttcttgctgaaaaaacattagttaaccaagcatcaaccaaaccagaagttcaagaagcc
 40  ttacaaaaagcaaaagaacttaatgaagctatgaaagcactgaaaactgaaataaataaa
     aagaacaaatcaaggctgatagtagatatgtaaatgctgacagtggtcttcaagcaaat
     tacaattctgcgttaaattatggttctcaaattattgcaactacccaaccaccagagctt
     aataaagatgtaataaaatagagcaactcaaacgattaaaactgctgaaaataatttaaat
     gggcaatctaaattagcagaggctaagtcagacggaaatcaaagcatcgaacatttgcaa
 45  ggattaacacaatcacaaaaagataaacaacatgatttaattaatcaagctcaaactaaa
     caacaggtagatgatatcgtaaataactctaaacaattagataactctatgaatcaacta
     caacaaattgttaacaatgacaatacagtaaaacaaaatagtgatttcattaatgəagat
     tccaʰccʰacaggatgcttataatcatgcaattcaagcagcaaaagatttgataaˈtgct
     catcctactatcatggataaaaatcaaatagatcaagctattgaaaatatcaaacaagca
 50  cttaatgatttacacggtagtaataaactatcagaagataaaaaagaagcttcagaacaa
     ctacaaaaccttaatagcttgacgaacgggcaaaaagatacgattttaaatcatattttc
     agtgcaccaacaagaagccaagtaggagaaaaaattgcaagtgctaaacaattaaataat
     acaatgaaagcacttagagattctattgctgataataatgaaatttttacaaagtagtaag
     tacttcaatgaagattctgaacaacaaaatgcttataatcaagccgtaaataaagctaaa
 55  aatataattaatgatcaaccaacaccagtaatggcaaatgatgagattcaaagtgtccta
     aatgaagttaaacaaactaaagataatttacatggtgatcaaaaacttgctaacgacaag
     acagatgcccaagcaacattaaatgcgttaaattacttaaatcaagcgcaaagaggtaat
     cttgaaactaaagttcaaaactctaattctagaccagaagtacaaaaagtagttcaatta
     gcaaatcàacttaatgatgcgatgaaaaaattagatgatgctttaactggtaatgacgca
```

                                    733

```
        ataaaacaaacgagtaattatattaatgaagatacttctcaacaagttaactttgatgag
        tatacagatagaggtaaaaacatagttgctgaacaaacaaatccaaatatgtctccaact
        aatattaacactattgctgataaaattactgaagctaaaaatgatttacatggcgtacaa
        aaactagaacaagctcaacaacagtccatcaatactattaatcaaatgactggtctaaac
5       caagctcaaaaagaacaattaaatcaagaaattcaacaaactcaaacccgttctgaagta
        catcaagtaattaataaagcacaagcttttaaatgattcaatgaatactttacgtcaaagt
        attactgatgaacatgaagttaaacaaacaagtaactacatcaatgaaactgttggtaat
        caaactgcatataacaatgccgttgatcgtgtaaaacaaataatcaatcaaacatctaat
        ccaactatgaatcctttagaggtggaacgtgcaacatcaaatgtaaaaacttctaaagat
10      gcacttcatggtgaacgtgaattgaatgacaataaaaattcaaaaacttttgcagtcaat
        cacttagataacctcaatcaagctcaaaaagaagcattaactcatgaaattgaacaagca
        actatagtttcacaagtaaataatatctataacaaagcgaaagctttaaataatgatatg
        aaaaaacttaaagatatcgttgctcaacaagataatgtgagacaatcaaacaattatata
        aacgaggatagtacacctcaaaatatgtacaacgatacaattaatcatgcacaatcaatc
15      attgatcaagtagcaaaccctacgatgtctcatgacgaaatagagaatgcaatcaataac
        ataaagcatgccatcaatgcactcgatggagaacataaattacaacaagcaaaagaaat
        gcaaacttattgattaatagtttaaacgatttaaatgcaccacaaagagatgccataaat
        agattggttaatgaagctcaaacaagagaaaaagtagctgaacaacttcaaagtgctcaa
        gctttaaatgacgctatgaagcatttaagaaacagcattcaaaatcaatcatccgtaaga
20      caagagagcaaatatattaatgcaagtgatgctaaaaaagagcaatataatcacgcagtt
        agagaagtcgaaaatattatcaatgaacaacatccaacattggataaagaaataattaag
        caactaacggatgctgtaaatcaagcgaataatgacttaaatggcgttgaattattagat
        gctgataagcaaaacgcacatcaatcgatacctacattgatgcacttaaatcaagcacaa
        caaaacgcattaaatgaaaaaattaataacgcagttaccagagctaaagttgcggctatt
25      attggccaagcaaaaatactcgatcatgctatggagaatttagaagaaagtatcaaagat
        aaagagcaagtcaaacagtcaagtaactatattaatgaagaccctgatgttcaagaaaca
        tacaataacgccgttgatcatgtgacagaaatacttaatcaaacagtaaatccaacttta
        tctattgaagatatagagcatgctatcaacgaagttaatcaagcgaaaaaacaactcaga
        ggtaaacaaaaactttatcaaactatcgatttagctgataaagaattaagtaaattggat
30      gatttaacatcacaacaaagcagttcaatatctaatcaaatatatactgctaaaacgaga
        acagaagttgcccaagcaattgaaaaagcaaaatcattaaatcatgcaatgaaagcactt
        aacaaaatatataaaaatgcagataaagtgttagatagtagtcgattcattaacgaagat
        caacctgaaaaagaggcgtatcaacaagctataaatcatgttgattcaatcattcataga
        caaacaaatcctgaaatggatccaacagtaatcaatagcataactcatgaactcgaaaca
35      gctcaaaataacttacatggtgatcagaaacttgctcatgcaaaacaagatgccgctaat
        gtaattaatggtctaattcatcttaatgttgctcaacgcgaggtaatgataaatacgaat
        acaaatgctacaacacgcgaaaaagttgcaaagaacttagataatgctcaagctcttgat
        aaagctatggaaacactacaacaagtagttgctcataaaaataatatattgaacgatagt
        aaatatttaaatgaagattcaaaatatcaacaacaatacgatcgagttattgctgatgcc
40      gaacaactacttaatcagcaacaaatccaacattagaaccttataaagtcgatattgtt
        aaggataatgtcctagctaacgaaaaaatactatttggcgcagaaaaactatcatatgac
        aaatcaaatgcaaatgatgaaattaaacatatgaattatcttaataatgcacaaaagcaa
        tctataaaagatatgatttctcacgcagcattaagaactgaagttaaacaacttctgcaa
        caagctaaaacccttgatgaagctatgaaatcacttgaagataaaactcaagtagtgatt
45      acagatactactttgcctaattacactgaagcttcagaggataaaaaggaaaaagtagac
        caaactgtatcacatgctcaagcaatcattgataaaataaatggctcaaatgtaagttta
        gatcaagtacgacaagcactagaacaattaactcaagcatcagaaaacctcgatggtgat
        cagcgagttgaagaagctaaagttcatgctaatcaaacaattgaccaattaacacatctt
        aattcattacaacaacaaactgcgaagaaagtgttaaaaacgcaacaaaactagaagaa
50      atcgctactgctagtaacaatgctctggcattaaacaaagtaatgggtaaattagaacaa
        ttcattaatcatgctgattctattgaaaatagtgataattatagacaagccgacgacgac
        aaaattatcgcttatgatgatgcactagaacatggacaagatatacaaaaatctaacgca
        acccaaaatgaagcaaaacaagcgttacaacaattaataaatgcagaaacatcgttaaat
        ggtttcgaaagattaaatcatgctagaccacgagctttagaatatattaaatcactagaa
55      aaataaacaatgctcaaaagtctgctttagaggataaagtaacgcaatcgcatgatttta
        ttagaattagaacatcttgtcaacgagggcacaaacctcaatgacattatgggtgaatta
        gctaacgcaatcgttaataactatgctccaaccaaagcaagtataaattatattaacgcc
        gataacctacgcaaagataactttactcaagctatcaacaatgcacgtgatgcactcaac
        aaaacucaaggtcagaacttagatttcaatgcaattgatacatttaaagatgatattattc
```

aaaactaaagatgcacttaacggtattgaacgtttaacagctgcaaaatcaaaagcagaa
aaactaattgatagtttaaaatttattaataaagctcaattcacacatgcaaatgatgaa
attatgaatactaattctattgcacaattgtctagaatcgtgaatcaagcatttgattta
aatgatgcaatgaaatctttaagagatgaacttaataatcaagcttttcctgtccaagca

5   agctcaaattatataaattcagatgaagatttaaaacaacaatttgaccatgctttaagt
aatgctcgaaaagtacttgcaaaagaaatggtaaaaatttagatgaaatacaaattgag
ggactcaaacaagtgattgaggatactaaagatgctttaaatggtatccaacgtttatca
aaagctaaagctaaagcaattcaatacgtacaatctttatcttatatcaatgatgcacag
cgtcatattgctgaaagtaatattcacaactctgatgatttatcatctttagcaaataca

10 ttatctaaagctagtgatttagataatgcaatgaaagacttacgagatactctagaaagt
aattcaacttctgttccaaatagtgtgaattatattaatgctgataagaatttacaaatt
gaatttgatgaggcgctacaacaagcaagtgcaacaagttctaaaacttcagaaaatcca
gcaacgattgaagaagtattaggtcttagtcaagccatttacgatacaaaaaatgcattg
aatggtgaacaacgtcttgcaactgagaagagcaaagatttaaaattaataaaaggatta

15 aaagatttaaataaagcacaacttgaagatgtcacaaacaaggtaaattcagcaaatact
ttaacagagttatctcagctcactcaatcaacgttaaaattaaacgataaaatgaaatta
ttgagagataagcttaaaaaccttagtaaatcctgttaaagcaagtttaaattatagaaac
gctgattataatttaaaacgtcaatttaacaaagctttaaaagaagctaaaggcgtatta
aataaaaatagcggtacaaatgtcaatatcaatgacattcaacatctttaacacaaata

20 gataatgctaaagaccaattaaatggtgaacgacgtctaaaagaacatcaacaaaatct
gaagtatttattattaaagaattagatatacttaataatgctcaaaaagctgcaataatt
aatcagattagagcgtctaaagacattaaaataattaatcaaatcgttgataatgcaata
gaattaaatgatgctatgcaaggtttaaaagaacatgtagctcaattaacagcaactaca
aaagacaacattgaatatttaaatgctgatgaagaccttaaaatacaatatgattacgct

25 atcaacttagcgaataatgttcttgacaaagaaaacggtacaaataaagacgctaatatc
ataattggaatgattcaaaacatggatgatgctagagcacttctaaatggaattgaaaga
cttaaagatgctcaaacaaaagcacataatgacattaaagatacgctcaaacgtcaactt
gatgaaattgaacacgctaatgcaacatcaaattctaaagctcaagcaaacaaatggta
aatgaggaagctagaaaagcgttttctaatattaatcacgcaacatcaaatgatttagtt

30 aatcaagcaaaagatgaagggcaatctgcaattgaacacatacatgcagatgaattacct
aaagcaaaactagatgctaatcaaatgattgaccaaaaagttgaagatataaatcacta
attagtcaaaatccaaacttatcaaatgaagaaaaaaataaactaatatctcaaattaat
aagttagtaaatggaattaagaatgaaattcaacaagctataaacaaacaacaaatagaa
aatgctacaacaaaactagatgaagtcattgaaactactaaaaaattaattatcgccaaa

35 gcagaagctaaacaagtgataaaagagttatcacaaaagaaacgagatgcaataaataac
aacactgatttaacacccttctcaaaaggcacatgctttagcagatattgataaaacagaa
aaagatgcacttcaacatatcgaaaattctaattcaattgatgatatcaataacaataaa
gagcatgcatttaatactttagctcatatcattatttgggatactgatcagcaaccatta
gtttttgaactacctgaattgagccttcaaaatgctctagtaacaagtgaggtggttgtt

40 cacagagatgaaactatttcattagaatctataattggagctatgactttaactgatgaa
cttaaagtcaatattgtttcattaccgaacactgataaagtagctgatcacctaaccgct
aaagttaaggttattttagctgatggctcatttgtcactgtaaatgttccagtcaaagtt
gtagaaaaagaattacaaatagctaaaaaggatgctataaaaacaattgatgttctggta
aaacaaaaatcaaagatatagattctaataacgaattaacgtctactcaacgtgaagat

45 gcaaaagctgaaattgaaagattgaaaaagcaagccatcgataaagtgactcattctaaa
tcgattaaagatattgaaacagtaaaacgaactgattttgaagaaatagatcagtttgat
cctaaacgctttacgctaaataaagctaaaaaggatatcattactgatgttaatactcaa
atccaaaatggtttcaaagaaattgaaacaataaaaggtttaacttctaatgaaaaaact
cagtttgataaacaattaactgcactacaaaaagaatttttagaaaaagtcgagcatgct

50 cataatttagtagaattaaatcaattacaacaagagtttaataatagatatgaacatatt
ttaaacccagcacattactaggtgaaaaacatatagcagaacataaattaggatctgtt
gtagtaaacaaaactcagcaaatactaaataatcaatctgcttcttactttataaaacaa
tgggcacttgatagaattaaacaaattcaactagaaacgatgaattcaattcgtggtgcg
cataccgtaca

55 Sequence 2964
MGTLKSLVAKQPTVQKTSVYINEDQPEQSAYNDSITMGQTIINKTADPVLDKTLVDNAIS
NISTKENALHGEQKLTTAKTEAINALNTLADLNTPQKEAIKTAINTAHTRTDVTAEQSKA
NQINSAMHTLRQNISDNESVTNESNYINAEPEKQHAFTEALNNAKEIVNEQQATLDANSI

```
         NQKAQAILTTKNALDGEEQLRRAKENADQEINTLNQLTDAQRNSEKGLVNSSQTRTEVAS
         QLAKAKELNKVMEQLNNLINGKNQMINSSKFINEDANQQQAYSNAIASAEVLKNKSQNPE
         LDKVJIEQAINNINSAINNLNGEAKLTKAKEDAVASINNLSGLTNEQKTKENQAVVGSQT
         RDQVANVLRDSKALDQSMQTLRDLVNNQNVIHSTSNYFNEDSTQKNTYDNAIDNGSTYIT
   5     GQHNSELNKSTIDQTISQINTAKNDLHGAEKLQRDKGTANQEIGQLGYLNDPQKSAEESL
         VNGSNTRSEVEEHLNEAKSLNNAMKQLRDKVAEKTNVKQSSDYINDSTEHQRGYDQALQE
         AENIINEIGNPTLNKSEIEQKLQQLTDAQNALQGSHLLEEAKNNAITEINKLTALNDAQR
         QKAIENVQAQQTIPAVNQQLTLDREINTAMQALRDKVGQQNNVHQQSNYFNEDEQPKHNY
         DNSVQAGQTIIDKLQDPIMNKNEIEQAINQINTTQTALSGENKLHTDQESTNRQIEGLSS
   10    LNTAQINAEKDLVNQAKTRTDVAQKLATAKEINSAMSNLRDGIQNKEDIKRSSAYINADP
         TKVTAYDQALQNAENIINATPNVELNKATIEQALSRVQQAQQDLDGVQQLANAKQQATQT
         VNGLNSLNDGQKRELNLLINSANTRTKVQEELNKATESNHAMEALRNSVQNVDQVKQSNN
         YVNEDQPEQHNYDNAVNEAQATINNNAQPVLDKLAIERLTQTVNTTKDALHGTQKLIQDQ
         QAAETGIRGLTSLNEPQKNAEVAKVTAATTRDEVRNIRQEATTLDTAMLGLRKSIKDKND
   15    TKNSSKYINEDHDQQQAYDNAVNNAQHVIDETQATLSSDTINQLANAVTQAKSNLHGDTK
         LQHDKDSAKQTIAQLQNLNSAQKHMEDSLIDNESTRTQVQHDLTEAQALDGLMGALKESI
         KDNTNIVSNGNYINAEPSKKQAYDAAVQNAQNIINGTNQPTINKGNVTTATQTVKNTKDA
         LDGDHRLEEAKNNANQTIRNLSNLNNAQKDAEKNLVNSASTLEQVQQNLQTAQQLDNAMG
         ELRQSIANKDQVKADSKYLNEDPQIKQNYDDAVQRVETIINETQNPELLKANIDQATQSV
   20    QNAEQALHGAEKLNQDKQTSSTELDGLTDLTDAQREKLREQINTSNSRDDIKQKIEQAKA
         LNDAMKKLKEQVAQKDGVHANSDYTNEDSAQKDAYNNALKQAEDIINNSSNPNLNAQDIT
         NALNNIKQAQDNLHGAQKLQQDKNTTNQAIGNLNHLNQPQKDALIQAINGATSRDQVAEK
         LKEAEALDEAMKQLEDQVNQDDQISNSSPFINEDSDKQKTYNDKIQAAKEIINQTJNPTL
         DKQKIADTLQNIKDAVNNLHGDQKLAQSKQDANNQLNHLDDLTEEQKNHFKPLINNADTR
   25    DEVNKQLEIAKQLNGDMSTLHKVINDKDQIQHLSNYINADNDKKQNYDNAIKEAEDLIHN
         HPDTLDHKALQDLLNKIDQAHNELNGESRFKQALDNALNDIDSLNSLNVPQRQTVKDNIN
         HVTTLESLAQELQKAKELNDAMKAMRDSIMNQEQIRKNSNYTNEDLAQQNAYNHAVDNIN
         NIIGEDNATMDPQIIKQATQDINTAINGLNGDQKLQDAKTDAKQQITNFTGLTEPQKQAL
         ENIINQQTSRANVAKQLSHAKFLNGKMEELKVAVAKASLVRQNSNYINEDVSEKEAYEQA
   30    IAKGQEIINSENNPTISSTDINRTIQEINDAEQNLHGENKLRQAQEIAKNEIQNLDGLNS
         AQITKLIQDIGRTTTKPAVTQKLEEAKAINQAMQQLKQSIADKDATLNSSNYLNEDSEKK
         LAYDNAVSQAEQLINQLNDPTMDISNIQAITQKVIQAKDSLHGANKLAQNQADSNLIINQ
         STNLNDKQKQALNDLINHAQTKQQVAEIIAQANKLNNEMGTLKTLVEEQSNVHQQSKYIN
         EDPQVQNIYNDSIQKGREILNGTTDDVLNNNKIADAIQNIHLTKNDLHGDQKLQKAQQDA
   35    TNELNYLTNLNNSQRQSEHDEINSAPSRTEVSNDLNHAKALNEAMRQLENEVALENSVKK
         LSDFINEDEAAQNEYSNALQKAKDIINGVPSSTLDKATIEDALLELQNARESLHGEQKLQ
         EAKNQAIAEIDNLQALNPGQVLAEKTLVNQASTKPEVQEALQKAKELNEAMKALKTEINK
         KEQIKADSRYVNADSGLQANYNSALNYGSQIIATTQPPELNKDVINRATQTIKTAENNLN
         GQSKLAEAKSDGNQSIEHLQGLTQSQKDKQHDLINQAQTKQQVDDIVNNSKQLDNSMNQL
   40    QQIVNNDNTVKQNSDFINEDSSQQDAYNHAIQAAKDLITAHPTIMDKNQIDQAIENIKQA
         LNDLHGSNKLSEDKKEASEQLQNLNSLTNGQKDTILNHIFSAPTRSQVGEKIASAKQLNN
         TMKALRDSIADNNEILQSSKYFNEDSEQQNAYNQAVNKAKNIINDQPTPVMANDEEQSVL
         NEVKQTKDNLHGDQKLANDKTDAQATLNALNYLNQAQRGNLETKVQNSNSRPEVQKVVQL
         ANQLNDAMKKLDDALTGNDAIKQTSNYINEDTSQQVNFDEYTDRGKNIVAEQTNPNMSPT
   45    NINTIADKITEAKNDLHGVQKLEQAQQQSINTINQMTGLNQAQKEQLNQEIQQTQTRSEV
         HQVINKAQALNDSMNTLRQSITDEHEVKQTSNYINETVGNQTAYNNAVDRVKQIINQTSN
         PTMNPLEVERATSNVKTSKDALHGERELNDNKNSKTFAVNHLDNLNQAQKEALTHEIEQA
         TIVSQVNNIYNKAKALNNDMKKLKDIVAQQDNVRQSNNYINEDSTPQNMYNDTINHAQSI
         IDQVANPTMSHDEIENAINNIKHAINALDGEHKLQQAKENANLLINSLNDLNAPQRDAIN
   50    RLVNEAQTREKVAEQLQSAQALNDAMKHLRNSIQNQSSVRQESKYINASDAKKEQYNHAV
         REVENIINEQHPTLDKEIIKQLTDAVNQANNDLNGVELLDADKQNAHQSIPTLMHLNQAQ
         QNALNEKINNAVTRAKVAAIIGQAKILDHAMENLEESIKDKEQVKQSSNYINEDPDVQET
         YNNAVDHVTEILNQTVNPTLSIEDIEHAINEVNQAKKQLRGKQKLYQTIDLADKELSKLD
         DLTSQQSSSISNQIYTAKTRTEVAQAIEKAKSLNHAMKALNKIYKNADKVLDSSRFINED
   55    QPEKEAYQQAINHVDSIIHRQTNPEMDPTVINSITHELETAQNNLHGDQKLAHAKQDAAN
         VINGLIHLNVAQREVMINTNTNATTREKVAKNLDNAQALDKAMETLQQVVAHKNNILNDS
         KYLNEDSKYQQQYDRVIADAEQLLNQTTNPTLEPYKVDIVKDNVLANEKILFGAEKLSYD
         KSNANDEIKHMNYLNNAQKQSIKDMISHAALRTEVKQLLQQAKTLDEAMKSLEDKTQVVI
         TDTTLPNYTEASEDKKEKVDQTVSHAQAIIDKINGSNVSLDQVRQALEQLTQASENLDGD
```

```
QRVEEAKVHANQTIDQLTHLNSLQQQTAKESVKNATKLEEIATASNNALALNKVMGKLEQ
FINHADSIENSDNYRQADDDKIIAYDDALEHGQDIQKSNATQNEAKQALQQLINAETSLN
GFERLNHARPRALEYIKSLEKINNAQKSALEDKVTQSHDLLELEHLVNEGTNLNDIMGEL
ANAIVNNYAPTKASINYINADNLRKDNFTQAINNARDALNKTQGQNLDFNAIDTFKDDIF
```
5
```
KTKDALNGIERLTAAKSKAEKLIDSLKFINKAQFTHANDEIMNTNSIAQLSRIVNQAFDL
NDAMKSLRDELNNQAFPVQASSNYINSDEDLKQQFDHALSNARKVLAKENGKNLDEIQIE
GLKQVIEDTKDALNGIQRLSKAKAKAIQYVQSLSYINDAQRHIAESNIHNSDDLSSLANT
LSKASDLDNAMKDLRDTLESNSTSVPNSVNYINADKNLQIEFDEALQQASATSSKTSENP
ATIEEVLGLSQAIYDTKNALNGEQRLATEKSKDLKLIKGLKDLNKAQLEDVTNKVNSANT
```
10
```
LTELSQLTQSTLKLNDKMKLLRDKLKTLVNPVKASLNYRNADYNLKRQFNKALKEAKGVL
NKNSGTNVNINDIQHLLTQIDNAKDQLNGERRLKEHQQKSEVFIIKELDILNNAQKAAII
NQIRASKDIKIINQIVDNAIELNDAMQGLKEHVAQLTATTKDNIEYLNADEDLKIQYDYA
INLANNVLDKENGTNKDANIIIGMIQNMDDARALLNGIERLKDAQTKAHNDIKDTLKRQL
DEIEHANATSNSKAQAKQMVNEEARKAFSNINHATSNDLVNQAKDEGQSAIEHIHADELP
```
15
```
KAKLDANQMIDQKVEDINHLISQNPNLSNEEKNKLISQINKLVNGIKNEIQQAINKQQIE
NATTKLDEVIETTKKLIIAKAEAKQVIKELSQKKRDAINNNTDLTPSQKAHALADIDKTE
KDALQHIENSNSIDDINNNKEHAFNTLAHIIIWDTDQQPLVFELPELSLQNALVTSEVVV
HRDETISLESIIGAMTLTDELKVNIVSLPNTDKVADHLTAKVKVILADGSFVTVNVPVKV
VEKELQIAKKDAIKTIDVLVKQKIKDIDSNNELTSTQREDAKAEIERLKKQAIDKVTHSK
```
20
```
SIKDIETVKRTDFEEIDQFDPKRFTLNKAKKDIITDVNTQIQNGFKEIETIKGLTSNEKT
QFDKQLTALQKEFLEKVEHAHNLVELNQLQQEFNNRYEHILNQAHLLGEKHIAEHKLGYV
VVNKTQQILNNQSASYFIKQWALDRIKQIQLETMNSIRGAHTVX
```

```
Sequence 2965
```
25
```
Contig_0541_pos_1894_2919
is similar to (with p-value 2.0e-28)
 >gp:gp|AB001577|AB001577_1 Pseudomonas sp. DNA for low spec
ificity L-threonine aldolase, complete cds. NID: g2865133.
gtgatttcatttgaaaatgattatttagaaggtgcacatgaaaaagttttaaatcgatta
```
30
```
gtagagacaaatcgaatacaagctgctggatatggcttcgatgacttttcggcacaagct
gcagataaaattagacaacgtattgactgtccagatgctaccattcgttttttagtaggt
ggtacgcaaaccaatcaagtagttattaactcaatgcttgatagttatgaaggtgttata
tccgctgatacaggacatgtggcagtccatgaaggtggtgcgatagaattcagtggacat
aaagctctaaccataccctcccaagaaggtaagattactgctcaagacgttgagacttat
```
35
```
atagaaacttttgaaagtgattttaaaaaagaacacatggtgtatccagggatggtttat
atttcacatccaaccgaatatggaactttatacacgaagaagaattacaatctttatct
agagtttgccgtagacatcagattccactatttatggatggtgcacgtttaggctatggc
cttatgagcaatcaaactaatgtaactatcgaagatgttgcaaaatactgtgatgtgttt
tacataggaggtactaagattggagcactttgtggtgaagcaattgtcttcactaaacaa
```
40
```
aatgaacctaaaaacttcactacaattataaaacatcatggtgctttattagcaaaaggc
cgtctaactggtgttcaatttttagaattattcactgatgatttatattttgatataagt
cgacatgctattaaaatggctgaaaaggtaaaaaaaggatttatagataaaggatatcaa
gtctattttgattcaccaaccaatcaacaattttttattttaagcaacgataaaattgaa
gaactaaaacaaaaggtaaaattcgcagtttgggagaaatacgataatcaacatcgtgta
```
45
```
gttcgcttcgcaacaagttgggccacaactgaagaaaatgttaatcaactacttgaacta
atataa
```

```
Sequence 2966
VISFENDYLEGAHEKVLNRLVETNRIQAAGYGFDDFSAQAADKIRQRIDCPDATIRFLVG
```
50
```
GTQTNQVVINSMLDSYEGVISADTGHVAVHEGGAIEFSGHKVLTIPSQEGKITAQDVENY
IETFESDFKKEHMVYPGMVYISHPTEYGTLYTKEELQSLSRVCRRHQIPLFMDGARLGYG
LMSNQTNVTIEDVAKYCDVFYIGGTKIGALCGEAIVFTKQNEPKNFTTIIKHHGALLAKG
RLTGVQFLELFTDDLYFDISRHAIKMAEKVKKGFIDKGYQVYFDSPTNQQFFILSKDKIE
ELKQKVEFAVWEKYDNQHRVVRFATSWATTEENVNQLLELI*
```
55
```
Sequence 2967
Contig_0541_pos_7422_6532
is similar to (with p-value 3.0e-74)
 >sp:sp|P54204|FUR_STAEP FERRIC UPTAKE REGULATION PROTEIN. >
```

gp:gp|X97011|SEFURSOD_1 S.epidermidis genes fur and sod (par
tial). NID: g1263907.
atgttaataggatcacatgtttcaatgagtggcaaaaaaatgctgcaagggtcagcagaa
gaagcacataaatatggtgaatctacatttatgatttatacaggtgcgcctcaaaataca
agacgtaaaaatattgaagatttaaatatcgaaaaaggccagcaggcaatgaaaacaatat
ggcttatcaaatatcgttgtacatgcaccatatatcattaacattgcaaatacaaccaaa
cctgaagtatttaatttaggagtcgactttctacaaaaagaaatcgaaagaactcaagcg
ctcggagcgaaagatattgtactgcatcctggagcgcatgtcggagcaggtgtagataaa
ggaattcaaaaaattattgaaggacttaatgaagtactcacacatgataatgatgtaaga
atagcacttgaaactatggcgggtaaaggaacagaagtagggagatcttttgaagaaatt
gctcaaataattgatggtgttacacataatgatcgcttatcagtatgttttgatacgtgc
cacactcatgatgccggttataacgtcaaagaagatttcgatggtgtactagaaaaattc
gacagcataattggagtagatcgaattaaagtagtacatgttaatgacagtaaaaaccta
agaggtgcacagaaggatcgtcacgaaaatatcggctttggtcatattggctttgatgca
cttaattacgtagtacatcatgatacttttaaaaatattcccaaaatattagaaactcca
tatgttggtgaagataaaaaaaataaaaaaccaccgtataaattagaaatagacatgtta
aaatcacaaaaatttgatccagaactcaaaaacaaaattttaactcaataa

Sequence 2968
MLIGSHVSMSGKKMLQGSAEEAHKYGESTFMIYTGAPQNTRRKNIEDLNIEKGQQAMKTY
GLSNIVVHAPYIINIANTTKPEVFNLGVDFLQKEIERTQALGAKDIVLHPGAHVGAGVDK
GIQKIIEGLNEVLTHDNDVRIALETMAGKGTEVGRSFEEIAQIIDGVTHNDRLSVCFDTC
HTHDAGYNVKEDFDGVLEKFDSIIGVDRIKVVHVNDSKNLRGAQKDRHENIGFGHIGFDA
LNYVVHHDTFKNIPKILETPYVGEDKKNKKPPYKLEIDMLKSQKFDPELKNKILTQ*

Sequence 2969
Contig_0541_pos_4647_4228
 >sp:sp|P54476|END4_BACSU PROBABLE ENDONUCLEASE IV (EC 3.1.2
1.2) (ENDODEOXYRIBONUCLEASE IV). >gp:gp|D84432|BACJH642_144
Bacillus subtilis DNA, 283 Kb region containing skin element
. NID: g2627063. >gp:gp|Z99116|BSUB0013_223 Bacillus subtili
s complete genome (section 13 of 21): from 2395261 to 261373
0. NID: g2634723.
atgaatacaaatgatgcaattaaagtttttaaaggaaaacggacttaaatatactgataaa
cgtaaagatatgctagatatctttgttaaagaggataaatatttaaatgctaaacatatt
caacaacaaatggataaagactatcctggaatatcatttgatactgtatacagaaatctt
catttatttaaagatttaggcattatagagagtaccgaattagatggagaaatgaaattc
agaatcgcatgcacaaatcatcaccatcatcattttatttgcgaaaattgcggagaaact
aaagtgattgattttgtccaatagaaaagattaaaagtcaattacccaatgtaaatatt
catactcataaattagaagtgtatggtatttgtgaagaatgtcaacgtaaagcaaactaa

Sequence 2970
MNTNDAIKVLKENGLKYTDKRKDMLDIFVKEDKYLNAKHIQQQMDKDYPGISFDTVYRNL
HLFKDLGIIESTELDGEMKFRIACTNHHHHHFICENCGETKVIDFCPIEKIKSQLPNVNI
HTHKLEVYGICEECQRKAN*

Sequence 2971
Contig_0542_pos_5747_5274
is similar to (with p-value 2.0e-17)
 >sp:sp|P41893|PPAL_SCHPO LOW MOLECULAR WEIGHT PHOSPHOTYROSI
NE PROTEIN PHOSPHATASE (EC 3.1.3.48) (LOW MOLECULAR WEIGHT C
YTOSOLIC ACID PHOSPHATASE) (EC 3.1.3.2) (PTPASE) (SMALL TYRO
SINE PHOSPHATASE). >pir:pir|A55446|A55446 protein-tyrosine-p
hosphatase (EC 3.1.3.48), low molecular weight - fission yea
st (Schizosaccharomyces pombe) >gp:gp|L33929|YSPLMPTP_1 Schi
zosaccharomyces pombe low Mr protein tyrosine phosphatase mR
NA, complete cds. NID: g602991.
gtgatactaatgatacatgtagcatttgtatgtctcggtaatatatgtcgttctccaatg
gctgaggctatcatgagacaaagactacaagaaagaggtatttcagatataaaagttcat

```
tctagaggaacaggacgttggaatttaggcgaacctccacataacggaacacaaaaaatt
ctacagaagtaccatattccttatgatggtatggtgagtgaacttttcgaacctgatgat
gattttgactatattattgctatggaccaaagtaacgtagacaatatcaaacaaatcaat
ccaaatttacaaggacaattgttcaaattgctagaatttagtaacatggaagagagtgat
```
5
```
gtaccagatccatactacacaaataattttgaaggtgttttcgagatggtgcaatcatct
tgtgataaatttaatagactacatcgtaaaagatgcaaatttgaaagagaggtaa
```

Sequence 2972
VILMIHVAFVCLGNICRSPMAEAIMRQRLQERGISDIKVHSRGTGRWNLGEPPHNGTQKI
10   LQKYHIPYDGMVSELFEPDDDFDYIIAMDQSNVDNIKQINPNLQGQLFKLLEFSNMEESD
VPDPYYTNNFEGVFEMVQSSCDNLIDYIVKDANLKER*

Sequence 2973
Contig_0546_pos_1878_160
15   >sp:sp|P17444|BETA_ECOLI CHOLINE DEHYDROGENASE (EC 1.1.99.1
) (CHD). >gp:gp|X52905|ECBET_5 Escherichia coli betT, betI,
betB and betA genes. NID: g48714. >gp:gp|M77738|ECOBETA_1 E.
coli choline dehydrogenase (betA) gene, complete cds. NID: g
145401. >gp:gp|AE000138|AE000138_2 Escherichia coli K-12 MG1
20   655 section 28 of 400 of the complete genome. NID: g1786501.

```
atgagaagaaaacgcgattcatacgattatgtcatcattggtggcggtagtgcaggttca
gttcttggtgcacgcctttcagaggataaagataaaaatgttttggtattagaagctgga
cgtagtgactatttctgggatttatttattcaaatgccagcagcattgatgttcccatca
```
25
```
ggtaatcgttttttatgactgggaatatcaaactgacgaagaaccacatatgggacgtaga
gtagatcatgcgagaggtaaagtattaggtggctcaagttctattaacggtatgatttat
caacgaggtaacccaatggactatgaaggatgggcagaacctgaaggaatggacacatgg
gactttgcacattgtctaccatacttcaaaaagttagaaacaacatatggtgcagcgcca
tacgataaagttagaggccatgatggtccaatcaaattaaaacgtggaccagctactaat
```
30
```
ccattatttaaatcattctttaatgcaggtgttgaagcgggctatcataaaactgcagac
gttaatggatacagacaagaaggttttggaccatttgatagccaagtacatcatggacgt
cgtatgtctgcttcaagagcgtatctacgcccagcattaagacgtagaaacttagatgtt
gaaacacgtgcattcgttacaaaattaattttttgatgaaaataatagtaaaaaagtaaca
ggcgtgactttcaagaaaaatggtaaagaacatactgttcatgcaaacgaagttatttta
```
35
```
tctggcggtgctttcaatacaccacaactattacaattatcaggtattggtgactcagaa
ttcttaaaatcaaaaggtatagagccacgtatgcatttaccaggtgttggtgagaacttc
gaagatcacttagaagtatatattcaacataaatgtaaacaaccggtttcactacaacct
agccttgatgtcaaacgtatgccgttcatcggtttacaatggattttttgcacgtaaaggt
gcagcggcgtctaaccactttgaaggtggtggctttgtaagatcaaatgatgatgttgat
```
40
```
tatccaaacctcatgttccatttcttaccaattgctgtaagatatgatggtcaaaaagca
ccagtagcacatggttaccaagtacatgttggaccaatgtactccaactcaagaggtagt
ttgaaaatcaaatctaaggatccatttgaaaaaccaagtatcgtgtttaattacttatct
acgaagaagacgaaagagaatgggttgaagcaattagagtagcaagaaatatcctaaaa
caaaaagctatggacccatttaatggtggcgaaatttcaccaggaccacaagttcaaacg
```
45
```
gatgaagaaattctagattgggtacgtaaagatggagaaactgcattacatccatcttgt
agcgcgaaaatgggacctgcatctgacccaatggcagtagtcgatccattaactatgaaa
gtacatggtatggaaaatttacgtgtcgttgatgcttcagcaatgcctagaacaacaaat
ggtaatattcatgcacctgtattgatgttagctgagaaagcagcggacattattcgtggt
agaaaaccgcttgaacctcaatatgttgactattataaacatggtattgatgatgaaaaa
```
50
```
gcaggtgcaatggaagatgatccattctaccaatattaa
```

Sequence 2974
MRRKRDSYDYVIIGGGSAGSVLGARLSEDKDKNVLVLEAGRSDYFWDLFIQMPAALMFPS
GNRFYDWEYQTDEEPHMGRRVDHARGKVLGGSSSINGMIYQRGNPMDYEGWAEPEGMDTW
55   DFAHCLPYFKKLETTYGAAPYDKVRGHDGPIKLKRGPATNPLFKSFFNAGVEAGYHKTAD
VNCYRQESFGPFDSQVHHGRRMSASRAYLRPALRRRNLDVETRAFVTKLIFDENNCKKVT
GVTFKKHGKEHTVHANEVILSGGAFNTPQLLQLSGIGDSEFLKSKGIEPRMHLPGVGENF
EDHLEVYIQHKCKQPVSLQPSLDVKRMPFIGLQWIFARKGAAASNHFEGGGFVRSNDDVD
YPNLMFHFLPIAVRYDGQKAPVAHGYQVHVGPMYSNSRGSLKIKSKDPFEKPSIVFNYLS
```

TKEDEREWVEAIRVARNILKQKAMDPFNGGEISPGPQVQTDEEILDWVRKDGETALHPSC
SAKMGPASDPMAVVDPLTMKVHGMENLRVVDASAMPRTTNGNIHAPVLMLAEKAADIIRG
RKPLEPQYVDYYKHGIDDEKAGAMEDDPFYQY*

5   Sequence 2975
       Contig_0547_pos_6940_6371
       is similar to (with p-value 2.0e-34)
         >pir:pir|S47148|S47148 hypothetical protein 1 - Staphylococ
       cus carnosus >gp:gp|X79725|SCSECA_1 S.carnosus (TM300) secA
10   gene. NID: g499333.
       atgattagatttgaaattcatggagataacctcactatcacagatgcaattcgcaactat
       attgaggagaaagtaggtaaattagaaagatactttaacaatgtgccaaatgctgtagca
       catgttagagttaaaacttattctaattctacaactaaaattgaagttacaattccttta
       aaagatgtcactcttagagctgaagaaagacatgatgatttatatgcaggcattgattta
15   attactaacaaattggaaagacaagtacgtaaatacaaaactcgtgtaaatcgtaaacat
       agagatcgtggagatcaagatatctttgttgctgaagtacaagagtctacaacaaacaat
       catgcagatgatatagaaagtgaaaatgatattgaaattattcgttctaaacaattcagc
       ttaaaaccaatggattctgaagaagcagtattgcaaatgaatttattgggacatgatttc
       tttattttaccgatagagaaactgatggcacaagcattgtttatagacgaaaagatggc
20   aaatacggtctgattgaaacaactgaataa

       Sequence 2976
       MIRFEIHGDNLTITDAIRNYIEEKVGKLERYFNNVPNAVAHVRVKTYSNSTTKIEVTIPL
       KDVTLRAEERHDDLYAGIDLITNKLERQVRKYKTRVNRKHRDRGDQDIFVAEVQESTTNN
25   HADDIESENDIEIIRSKQFSLKPMDSEEAVLQMNLLGHDFFIFTDRETDGTSIVYRRKDG
       KYGLIETTE*

       Sequence 2977
       Contig_0550_pos_2797_2369
30   is similar to (with p-value 5.0e-26)
         >gp:gp|AF036166|AF036166_1 Xanthomonas campestris organic h
       ydroperoxide resistance protein (ohr) gene, complete cds. NI
       D: g3098341.
       atggcaaattctatttactcaactacaatgatttcaaatggtggacgtgatggtcgcgtt
35   tttagtccagataatacatttgttcaaaaccttgcaacacctaaggaaatgggtggtcaa
       ggaggcaacgatactaatcctgaacaattatttgctgctggatatagtgcatgctttaat
       agcgcgctatcattaatcttatctcaaaataaaataagtgatgccaacccagaagttgaa
       atcactattgaattacttaaagatgatactgacaatggttttaaacttggcgcagatatt
       aaagtcacacttgaaaatatgtcccaacaagatgctgagaaatttgtggagcaagcacat
40   caattctgtccatactcaaaagcgacacgtggtaacattgacgttcagttagatgttaca
       gcgcaataa

       Sequence 2978
       MANSIYSTTMISNGGRDGRVFSPDNTFVQNLATPKEMGGQGGNDTNPEQLFAAGYSACFN
45   SALSLILSQNKISDANPEVEITIELLKDDTDNGFKLGADIKVTLENMSQQDAEKFVEQAH
       QFCPYSKATRGNIDVQLDVTAQ*

       Sequence 2979
       Contig_0553_pos_7975_7067
50   is similar to (with p-value 1.0e-37)
         >sp:sp|Q27546|IUNH_CRIFA INOSINE-URIDINE PREFERRING NUCLEOS
       IDE HYDROLASE (EC 3.2.2.1) (IU-NUCLEOSIDE HYDROLASE) (PURINE
         NUCLEOSIDASE).
       atgtctattcccattataattgatactgatcctggtatagatgatgctacagcaattagt
55   atcgcactttcacaccctcaatttgacgttaaaatgatatcaactgtgaatggtaatgta
       ggtattgagaaaacgacagcaaatgcattaaagctaaaaaggtttttaatagttctgtt
       cctgtacatagaggggcatcccaaccattgattaatgacatctttgaagctacatcaatt
       catggtgagtctggtatggatggttacgagtttccacaaataaatcaagatgatttaaca
       tcaattcatgcagttgaagcaatgagaaatctattagtaaatactcaagaacctttaacc

```
ttgattgccataggtccactaacaaatatcgctattctttttaactagttatcccgaagtt
caaccatttattaaggaaattgttttaatgggtggtagtaccggtagaggtaatgtaacg
cctttagctgaatttaatatatattgtgatccagaagctgctcaaattgtatttaactct
ggattacctttgactatgattggtcttgatttggctcgtgaagcattgtttactcaccat
```
5
```
tttgtaaaagacttcaaagatacaaatgcaacttcaaacatgttatataatttatttcag
cactataagagtgaagattttgaaataggatttaaattatacgatgtattcactatatta
tatttgttggatccagaagcctttaatgtcaaggaggcgtatactcaaatagaattaaat
ggcaactttacaaggggagccacagtggtagactttaatatggagcatcccaattgtaca
gttgttttaagtcctgttgaaagacagtatgaggatttattcttaaacgccctttcttat
```
10
```
tgtaaataa
```

Sequence 2980
MSIPIIIDTDPGIDDATAISIALSHPQFDVKMISTVNGNVGIEKTTANALKLKRFFNSSV
PVHRGASQPLINDIFEATSIHGESGMDGYEFPQINQDDLTSIHAVEAMRNLLVNTQEPLT
15
LIAIGPLTNIAILLTSYPEVQPFIKEIVLMGGSTGRGNVTPLAEFNIYCDPEAAQIVFNS
GLPLTMIGLDLAREALFTHHFVKDFKDTNATSNMLYNLFQHYKSEDFEIGFKLYDVFTIL
YLLDPEAFNVKEAYTQIELNGNFTRGATVVDFNMEHPNCTVVLSPVERQYEDLFLNALSY
CK*

20
Sequence 2981
Contig_0554_pos_5578_0
is similar to (with p-value 9.0e-26)
  >gp:gp|D87664|D87664_1 Thermus aquaticus DNA for DNA polyme
rase family X, aminopeptidase T, QAH/OAS sulfhydrylase, comp
25
lete cds. NID: g1526546.
```
atgacaaaaaaagatgtaattcaattattagaaaaaatagctatatatatggagctaaaa
ggagaaaatacatttaaagtttcagcgtatagaaaagccgcacaaagtctagaggttgat
gagcgtacattagaagagattgatgatgtaacagaacttaaaggcattggaaaaggcgta
ggagaagttattaatgaatttaaaacacaaggtcaatcatcgacccttcaagcacttcaa
```
30
```
gatgaagtacctgaagggttagtgccacttttgaaaatacaaggattaggcagtaaaaaa
atagcgaaactatatcatgaacttcaaattacagataaagaaatacttcaaaaagcctgt
gaagaaggtaaggtcagtgctttaaaaggttttgcaaaaaagacagagcaaaacatttta
gaagcagtgaagtcgatgggtgctaaaaaagatcgttatcctatagagctaatgagagga
ctcaaccaagaaattgtaaaatttattgaacagttagaaggagttgaacaatattcaact
```
35
```
gctggtagttttcgaagatataaggaaatgagtaaagatttagatttcataattagtaca
tcagagcctaaaaaagttcaacaacaattacttcgtattccgaataaagtcaaagatgtt
gctgttgggggatactaaaatttctctggaattagcttatgatgatgagacgattggcgtt
gattttagattgatagaacctt
```

40
Sequence 2982
MTKKDVIQLLEKIAIYMELKGENTFKVSAYRKAAQSLEVDERTLEEIDDVTELKGIGKGV
GEVINEFKTQGQSSTLQALQDEVPEGLVPLLKIQGLGSKKIAKLYHELQITDKEILQKAC
EEGKVSALKGFAKKTEQNILEAVKSMGAKKDRYPIELMRGLNQEIVKFIEQLEGVEQYST
AGSFRRYKEMSKDLDFIISTSEPKKVQQQLLRIPNKVKDVAVGDTKISLELAYDDETIGV
45
DFRLIEPX

Sequence 2983
Contig_0555_pos_3938_2001
  >gp:gp|D85230|PEEGLTD_4 Plectonema boryanum URF141, ORF243,
50
 NADH-dependent glutamate synthase large subunit (gltB) and
small subunit (gltD) and URF289 genes, partial and complete
cds. NID: g1339947.
```
atgagttacggctctatctcagcagaggcacatgagacgttggctcaagctatgaatcaa
attggaggtaaaagtaatagtggagaaggtggagaagattcttcacgttacgaaattcaa
```
55
```
aaggatggaagtaataagataagtgcgattaagcaagttgcatcaggtcgttttggggtg
acgagtgattacttgcaacatgcaaaagaaattcaaattaaagtcgcacaaggcgctaaa
ccaggggaaggtggacaactaccaggttcaaaagtatatccatggattgctgagactaga
ggttcgacaccaggtataggattaattttcaccaccaccacaccatgatatttattcaatt
gaggactcagcacagctcattcatgatttaaaaaatgcaaatagaagagctgatattgca
```

741

gttaagcttgtatcaaaaactggcgttggaactatagcttcaggggtagctaaagctttc
gccgataaaattgttataagtggttatgatggaggtacaggtgcatcgcctaaaacaagt
attcaacatgcaggtttgccatgggagataggccttgccgaaacacatcaaacacttaaa
ttaaatgatttgcgtagtcgcgtaaaattagaaacggatggtaagttactgacgggtaaa
5   gatgtagcttatgcttgtgcgcttggtgcagaagaatttggtttcgcaacagcaccactt
gttgtttggggtgtattatgatgagggtttgtcataacgatacgtgtccagtaggggtt
gcaacacaaacaaagatttaagagctttgtttagaggtaaggcacagcatgtagttaac
tttatgtattttatagctgaagaattacgtgaaattttggcttcacttggtttagaaaca
gtagaagagttagtaggaagaacagatcttcttcaacgttcgacgcaattgaaaccaaat
10   agtaaagcagcttcgcttcaaatagaacgtttaatagaacaatttgacgggggttaatacg
aaagagatatcacaaaccatcatcttgatgaaggattcgatttgaattatctgtaccca
gacgcacgctatagtattgaaaacgggcactcttttaccggaaattatgttgttaataat
gaacagcgagatgtaggtgtaattacaggtagtgcgatagctaaacaatatggagaagaa
ggattacctgaagatacgatacttgcttacactgaaggtcatgcaggtcaaagcttagct
15   gcatatgcaccacgcggattaacaatccatcataccggtgatgctaatgactacgtaggt
aaaggattgtccggtggaactgtcatcgtaaatgctccaaatagtcaacgtgaaaatgaa
attatagcaggaaatataaacttttacggggcttctagaggtaaagcgtttatcaatggt
aaagctggtgagcgtttctgtatcagaaatagtggtgcagatgttgtagtagaaggtatt
ggtgatcatggacttgaatatatgacaggggggacatgtcattatcttaggagatgttgga
20   aagaactttggccaaggcatgagcggggggcgtaagttatatttttctcttctgacgtggag
aaatttaaaaaggttaatgcgcttgaaactttagaattcagtagcatacgttttgatgag
gaaaaatctcttatcaaagacatgcttgaagcacattttaagcatacacgtagtaacaaa
gcacgccaattacttgaccaatttgacaatattgaaaagttagcaattaaagttattccg
aaagattacaaattaatgatgcaaaaaattgatttgaaaaaacgtcaaatggaacgtgaa
25   gatgaagcaacactggcagcgttttatgatgacagagaaacaattgaacaagagctacag
ccagcagtcatttattaa

Sequence 2984
MSYGSISAEAHETLAQAMNQIGGKSNSGEGGEDSSRYEIQKDGSNKISAIKQVASGRFGV
30   TSDYLQHAKEIQIKVAQGAKPGEGGQLPGSKVYPWIAETRGSTPGIGLISPPPHHDIYSI
EDLAQLIHDLKNANRRADIAVKLVSKTGVGTIASGVAKAFADKIVISGYDGGTGASPKTS
IQHAGLFWEIGLAETHQTLKLNDLRSRVKLETDGKLLTGKDVAYACALGAEEFGFATAPL
VVLGCIMMRVCHNDTCPVGVATQNKDLRALFRGKAQHVVNFMYFIAEELREILASLGLET
VEELVGRTDLLQRSTQLKPNSKAASLQIERLIEQFDGVNTKEISQNHHLDEGFDLNYLYP
35   DARYSIENGHSFTGNYVVNNEQRDVGVITGSAIAKQYGEEGLPEDTILAYTEGHAGQSLA
AYAPRGLTIHHTGDANDYVGKGLSGGTVIVNAPNSQRENEIIAGNINFYGASRGKAFING
KAGERFCIRNSGADVVVEGIGDHGLEYMTGGHVIILGDVGKNFGQGMSGGVSYIFSSDVE
KFKKVNALETLEFSSIRFDEEKSLIKDMLEAHFKHTRSNKARQLLDQFDNIEKLAIKVIP
KDYKLMMQKIDLKKRQMEREDEATLAAFYDDRETIEQELQPAVIY*

40

Sequence 2985
Contig_0555_pos_1962_520
  >sp:sp|P39812|GLTB_BACSU GLUTAMATE SYNTHASE [NADPH] LARGE C
HAIN (EC 1.4.1.13) (NADPH-GOGAT). >gp:gp|Z99113|BSUB0010_138
45   Bacillus subtilis complete genome (section 10 of 21): from
1781201 to 2014980. NID: g2634090. >gp:gp|Z99114|BSUB0011_9
Bacillus subtilis complete genome (section 11 of 21): from 2
000171 to 2207900. NID: g2634230.
atgaaatatgataaacagtcgctatcagaattgtctttggtagaccgtctttcgaatcat
50   gaagcgtttcaacaacgcttcactaaagaagatgcttcgattcaggtgcgcgctgtatg
gatcgcgggaacacctttttgtcaaactgggcaatcttatggaagagaaacaatagtgatgc
cctattggtaattatatacctgagtggaacgacttagtctatcatcaagattttaaagct
gcttacgaaagattgagagagacgaataattttcctgaatttacaggaagagtttgtcct
gcaccatgtgagcaatcatgtgttatgaaaattaatagagaatccgtggcgattaaaggt
55   attgaacgtacaattattgatgaagcatatgagaatgagtgggttcatcccgcatatcct
gaagatcataaagaccaacgagttgctatcgtaggtagtggtccagcgggacttacagca
gctgaagaattaaactttaaaggctataaagttactgtttatgaaaaggcgcatgaacca
ggcggcttgctaatgtatggtataccaaatatgaaactagataaagacgtaatacgtcga
cgtgtatcacttatgaaagatgctggggtttttatttaaaacaggcgttgaaattggcgtc

```
     gatgtgagccgtgaaacacttgaagaaaattatgatgctattattttatgcacaggtgct
     caaaatgcgagagatttaccattggaaggacgaatgggctctggtattcattttgcaatg
     gactatca.tactgaacaaacacagtatctaaatggtgagattgaaagtttgagca''tact
     gctaaagataagaatgtaattattataggtgctggtgatactggtgcagactgtgtagcg
5    acagcattacgtgaaaactgtaaatctattgttcaatttaataaatatacgaaacagcct
     gaagagattacttttgaaagtaatacttcctggccattagcaatgcctgttttcaaaatg
     gattatgcgcataaggaatatgaagctaaatttggtcaagaaccaagagcctatggtgta
     caaacaatgcgctatgatgttgacgagttaggaaatgttaaaggcttatatacacaaata
     ttaaaagaaacgcctgatggcatggtgatggaagatggaccagaacgattttggccggct
10   gatttagtcttattatctatagggtttgttggtactgaaaccactgttccgcatgcgttt
     gatatacacaccgagcgtaataaaattgtagctaatgatacaaattatcaaactaatcac
     gctaaaatatttgctgcaggagatgcaagacgaggtcagagtttggttgtttgggcaata
     aaagaaggtcgtgaagtagcacattctgttgatcaatacttaagtaaagaagttctagtg
     taa
15
     Sequence 2986
     MKYDKQSLSELSLVDRLSNHEAFQQRFTKEDASIQGARCMDCGTPFCQTGQSYGRETIGC
     PIGNYIPEWNDLVYHQDFKAAYERLRETNNFPEFTGRVCPAPCEQSCVMKINRESVAIKG
     IERTIIDEAYENEWVHPAYPEDHKDQRVAIVGSGPAGLTAAEELNFKGYKVTVYEKAHEP
20   GGLLMYGIPNMKLDKDVIRRRVSLMKDAGVLFKTGVEIGVDVSRETLEENYDAIILCTGA
     QNARDLPLEGRMGSGIHFAMDYLTEQTQYLNGEIESLSITAKDKNVIIGAGDTGADCVA
     TALRENCKSIVQFNKYTKQPEEITFESNTSWPLAMPVFKMDYAHKEYEAKFGQEPPAYGV
     QTMRYDVDELGNVKGLYTQILKETPDGMVMEDGPERFWPADLVLLSIGFVGTETTVPHAF
     DIHTERNKIVANDTNYQTNHAKIFAAGDARRGQSLVVWAIKEGREVAHSVDQYLSKEVLV
25   *


     Sequence 2987
     Contig_0557_pos_8699_9661
     is similar to (with p-value 1.0e-71)
30   >sp:sp|P76113|YNCB_ECOLI PUTATIVE NADP-DEPENDENT OXIDOREDUC
     TASE IN TEHB-RHSE INTERGENIC REGION (EC 1.-.-.-). >gp:gp|D90
     784|D90784_8 E.coli genomic DNA, Kohara clone #273(32.5-32.8
     min.). NID: g1742353. >gp:gp|D90785|D90785_2 E.coli genomic
     DNA, Kohara clone #274(32.7-33.0 min.). NID: g1742363.
35   atgccacaagacgatacatttaaatatgaagatatagatgttattgaaccttcagaaaat
     gaattgcaattgaaaacattatatatatcggttgatccatatatgagaggacgtatgaca
     aatgctgattcttatgtagatccgttcaaacaaggggaaccgttcaatggacatacggta
     tctaaagttttgaaatccaaggatagtaattttgatgaaggtgatatagtagtgggtatg
     ctaccttggagaaaaataaatacagtaaatagtgagtatgttaacaaagtacctacttct
40   gacgtaccattacatctttatcttagtgtgttggggatgcctggtcagactgcttatcat
     ggattacttgatattggacaacctaaagaaggtgaaacagtagttatttcagcagcttca
     ggt.gcagttggttcagttgtgggccaaattgcgaagcttaaaggttgcagagtcg''ggt
     atagctggtggagataaaaaagtgaactatctaaaaaatgaacttcgttttgatgctggt
     atcgattacaaaaaagataatttccctgaagcgttaaaagaagcggtgcctaacggtata
45   gatgtctacttcgaaaatgtaggtggatatattggcgatgaagtcttcaaacatctcaat
     acacatgcaagaattcctgtttgtggtgcgatttcatcctataatcatccagaaaaagat
     attggaccacgcattcagcaaacattgattaaaaatcaagcaatgatgagaggtttcata
     gtagcagaattcgctgatggttttaaagaagcgagcaaacaattagctcaatgggttcaa
     gagaataaaattaaaacacaagtttcagtagaagatggttttgataaagtgccgcaagcc
50   tttagaaatctgctaactggtgataattttggtaaacaagttattaaagtggcaagtgaa
     taa


     Sequence 2988
     MPQDDTFKYEDIDVIEPSENELQLKTLYISVDPYMRGRMTNADSYVDPFKQGEPFNGHTV
55   SKVLKSKDSNFDEGDIVVGMLPWRKINTVNSEYVNKVPTSDVPLHLYLSVLGMPGQTAYH
     GLLDIGQPKEGETVVISAASGAVGSVVGQIAKLKGCRVVGIAGGDKKVNYLKNELRFDAG
     IDYKKDNFPEALKEAVPNGIDVYFENVGGYIGDEVFKHLNTHARIPVCGAISSYNHPEKD
     IGPRIQQTLIKNQAMMRGFIVAEFADGFKEASKQLAQWVQENKIKTQVSVEDGFDKVPQA
     FRNLLTGDNFGKQVIKVASE*
```

Sequence 2989
Contig_0557_pos_9913_0
is similar to (with p-value 3.0e-17)
5   >sp:sp|Q08257|QOR_HUMAN QUINONE OXIDOREDUCTASE (EC 1.6.5.5)
      (NADPH:QUINONE REDUCTASE) (ZETA- CRYSTALLIN). >pir:pir|PNO4
      48|PNO448 zeta-crystallin / quinone reductase (NADPH) (EC 1.
      6.-.-) - human >gp:gp|L13278|HUMQUINZ_1 Homo sapiens zeta-cr
      ystallin/quinone reductase mRNA, complete cds. NID: g292414.
10
      atgccagttgacaaagcgccacgtgtacttggctttgatgctgttggtgtgattgaaaag
      ataggagatcaagtgtcaatgtttcaagaaggggacgtcgttttttattcaggttctcct
      aaccaaaatggttcgaatgaagaataccaattaatagaggaatatttagtagctaaagca
      cctacaaatttgaaaagtgaacaagcagctagcctacctttaactgggctaacagcttat
15   gaaacgctttttcgatgtttttggaatttcaaaagaaccatctgaaaataaaggtaaatca
      ttgttaataattaatggagcaggtggtgtaggtagtattgcaacacagatagcgaaattt
      tatgggttgaaggttattacaactgcttcgagagaggatactataaagtggtctgttaat
      atgggtgctgatgttgtactgaatcataagaaagatttaagtc

20   Sequence 2990
      MPVDKAPRVLGFDAVGVIEKIGDQVSMFQEGDVVFYSGSPNQNGSNEEYQLIEEYLVAKA
      PTNLKSEQAASLPLTGLTAYETLFDVFGISKEPSENKGKSLLIINGAGGVGSIATQIAKF
      YGLKVITTASREDTIKWSVNMGADVVLNHKKDLSX

25   Sequence 2991
      Contig_0558_pos_11954_0
      is similar to (with p-value 1.0e-22)
       >pir:pir|A43577|A43577 regulatory protein pfoR - Clostridiu
      m perfringens
30   gtgttatggggagtgtttgatatggatttgcttataggtacattattttttaattttagtt
      ttagttgttttttactctatttacatataaggcacctagtgggatgaaagcgatgggtgct
      ttagcgaatgcggcaattgcttcgtttttagttgaagctttcaataaatatgtgggtgga
      caagtatttggtattaaattcttagaagaattaggagatgctgctggaggtttaggtggt
      gtggctgcggctggattaactgcattagcgattggagtatctccagtttatgcactagtt
35   attggtgcagcttgtggaggtatggacttgttgccaggattttttcgctgggtatattgta
      ggctacatgatgaagtataccgagaaatatgtgccagatggtattgatttaattg

      Sequence 2992
      VLWGVFDMDLLIGTLFLILVLVVFTLFTYKAPSGMKAMGALANAAIASFLVEAFNKYVGG
40   QVFGIKFLEELGDAAGGLGGVAAAGLTALAIGVSPVYALVIGAACGGMDLLPGFFAGYIV
      GYMMKYTEKYVPDGIDLIX

      Sequence 2993
      Contig_0558_pos_7113_6400
45   is similar to (with p-value 2.0e-21)
       >pir:pir|A55345|A55345 diamine N-acetyltransferase (EC 2.3.
      1.57) - Escherichia coli >gp:gp|D25276|ECOSN1A_1 Escherichia
       coli gene for spermidine acetyltransferase, complete cds. N
      ID: g517104. >gp:gp|AE000254|AE000254_5 Escherichia coli K-1
50   2 MG1655 section 144 of 400 of the complete genome. NID: g17
      87862.
      atgagtcacttcgaacattttactttagaacattaccactcccttttttcataatgatcgg
      ttcatgtcggtgattttttaatacagttgcagtggcacttttatctgcatcgattgctaca
      gttataggtacatttggtgcgatcgctttatattacttacgcaacaaacggtttaaggtt
55   acgctattaacgatgaataatgtattaatggtatcttcagacgtagtcattggtgcttct
      tttttaattatgttcactgcaataggacatttcacaggtttaggttaggattttcaaca
      gtgctagcatctcatatagcgttctgtattccaattgttgttatcatcgtcttaccccaa
      ttatatgaaatgaatgataacatgttaaatgcagcaagagatttaggtgctaatgaatca
      caattattaactagcattattatacctaatattatgccctcgattataggaggattctttt

```
atggcattaacatattctctagatgattttacggtaagcttctttgttactggaaacgga
tttagtgtattgtctgtagaagtttatgctatggctcgaaaaggtataagtatggagatt
aatgccatttctacaataatattttattgctattatgtttggagtatttggatattacttt
attcaacatatcgttaatcgtcagaaaaagatgaagcgaggtgttaatgaatga
```

5

Sequence 2994
MSHFEHFTLEHYHSLFHNDRLMSVIFNTVAVALLSASIATVIGTFGAIALYYLRNKRFKV
TLLTMNNVLMVSSDVVIGASFLIMFTAIGHFTGLGLGFSTVLASHIAFCIPIVVIIVLPQ
LYEMNDNMLNAARDLGANESQLLTSIIIPNIMPSIIGGFFMALTYSLDDFTVSFFVTGNG
10   FSVLSVEVYAMARKGISMEINAISTIIFIAIMFGVFGYYFIQHIVNRQKKMKRGVNE*

Sequence 2995
Contig_0558_pos_1973_1638
is similar to (with p-value 1.0e-26)
15   >sp:sp|P23859|POTC_ECOLI SPERMIDINE/PUTRESCINE TRANSPORT SY
STEM PERMEASE PROTEIN POTC. >pir:pir|C40840|C40840 spermidin
e/putrescine transmembrane protein C - Escherichia coli >gp:
gp|D90747|D90747_3 Escherichia coli genomic DNA.(25.1 - 25.5
min). NID: g1651548. >gp:gp|AE000212|AE000212_10 Escherichi
20   a coli K-12 MG1655 section 102 of 400 of the complete genome
. NID: g1787358. >gp:gp|M64519|ECOPOTABCD_3 E.coli transport
protein (potA, potB, potC and potD) genes, complete cds. NI
D: g147325.
gtggtagagttattagaaattaactttatacatagaacttgtgaagtgttaattattatc
25   gatccgcagtatgcaaataatgggtacgcgaaaaaagcctttaaaatggctattgactat
gcttttttcagtattaaatatgaataaggtatacttatatgtggatattaagaatgagaaa
gcagtacatatctatcaaagtaataatttcgaaatagaaggaacgttaaaggaacacttc
tatacaaggggagaatatagagattgctatgtaatgggcttgttaaaaaggaattgggtt
aataagaatgatgatgatttgtctcatataagatga
30

Sequence 2996
VVELLEINFIHRTCEVLIIIDPQYANNGYAKKAFKMAIDYAFLVLNMNKVYLYVDIKNEK
AVHIYQSNNFEIEGTLKEHFYTRGEYRDCYVMGLLKRNWVNKNDDDLSHIR*

35   Sequence 2997
Contig_0561_pos_2326_3213
is similar to (with p-value 4.0e-30)
>sp:sp|P28246|BCR_ECOLI BICYCLOMYCIN RESISTANCE PROTEIN (SU
LFONAMIDE RESISTANCE PROTEIN).
40   atgatgatgactacaagttccaaacacttatctaaaatattaatcgttatacttggcgtt
atgactgcatttggtcctttgactattgatatgtacggaccatctttacctaaagttcag
catgcgtttggttcatcaatttcagaaatacaacttacattatcctttgctatgataggt
ttagctattggtcaatttgtattcgggccactatcagatgtattaggtcgtaaaaaaatg
gcactcattttattgattggatatttaatagcctcattattatcagtttttacagttcat
45   ttaacaatatttttaattatccgtttaattcaaggtttagcaggaggtggtgcaatcgtc
atagccaaagcttctattggagataactatgacggagacgaattagcaaaattttttaact
tctcttatggtcataaacggtataatcaccatcattgctccactgttaggtggcctcgct
ttatctattgcaagttggagaatgatttttatatttttaacaatcattaccttaatagtc
atcttaggcatttttattaaagatgccagttgggcccccatcaagaacaatctcagttaaat
50   tttaaagcaatatttaaagattttggtctgttattaacaaaacccaccttcgttattcca
atgttattgcaaggattaacttatgtcatgttattcagttattcgtcagccgcacctttt
atttcacaaaagatgtatcatatgacaccacttcaatacagtgcaatgtttgctattaat
ggagtgggtttgatagtcgtcagtcagataaccgctattatagtagaaaaggtaagccga
tatgcgatgctcatatatttaacaatcattcaaatgttaggtgtttaa
55

Sequence 2998
MMMTTSSKHLSKILIVILGVMTAFGPLTIDMYGPSLPKVQHAFGSSISEIQLTLSFAMIG
LAIGQFVFGPLSDVLGRKKMALILLIGYLIASLLSVFTVHLTIFLIIRLIQGLAGGGAIV
IAKASIGDNYDGDELAKFLTSLMVINGIITIIAPLLGGLALSIASWRMIFIFLTIITLIV
```

745

ILGILLKMPVGPHQEQSQLNFKAIFKDFGLLLTKPTFVIPMLLQGLTYVMLFSYSSAAPF
ISQKMYHMTPLQYSAMFAINGVGLIVVSQITAIIVEKVSRYAMLIYLTIIQMLGV*

Sequence 2999
Contig_0561_pos_4708_4139
is similar to (with p-value 3.0e-49)
  >sp:sp|P28368|YVYD_BACSU HYPOTHETICAL 22.0 KD PROTEIN IN FL
IT-SECA INTERGENIC REGION. >gp:gp|Z31376|BSFLIDST_5 B.subtil
is (HB2058) genes for FliD, FliS, FliT proteins. NID: g49937
9. >gp:gp|Z99122|BSUB0019_28 Bacillus subtilis complete geno
me (section 19 of 21): from 3597091 to 3809700. NID: g263602
9. >gp:gp|U56901|BSU56901_22 Bacillus subtilis putative tran
scriptional regulator (yvhJ), Ycr59c/YigZ homolog (yvhK), hi
stidine kinase (degS),transcriptional regulator of degradati
on enzyme (degU), (degV), (comFA), (comFB), (comFC), flagell
ar protein (yviB), negative regulator of flagellin (flgM), f
lagella protein (yviC), flagellar-hook associated protein 1
(flgK), flagellar-hook associated protein 3 (flgL), (yviE),
transmembrane protein (yviF), (csrA), flagellin (hag), flag
ellar protein (yviH), flagellar hook-associated protein 2 (f
liD), flagellar protein (fliS), flagellar protein (fliT), si
gma-54 modulator homolog (yviI), and (secA) genes, complete
cds. NID: g1762326.
atgattagatttgaaattcatggagataacctcactatcacagatgcaattcgcaactat
attgaggagaaagtaggtaaattagaaagatactttaacaatgtgccaaatgctgtagca
catgttagagttaaaacttattctaattctacaactaaaattgaagttacaattcctta
aaagatgtcactcttagagctgaagaaagacatgatgatttatatgcaggcattgattta
attactaacaaattggaaagacaagtacgtaaatacaaaactcgtgtaaatcgtaaacat
agagatcgtggagatcaagatatctttgttgctgaagtacaagagtctacaacaaacaat
catgcagatgatatagaaagtgaaaatgatattgaaattattcgttctaaacaattcagc
ttaaaaccaatggattctgaagaagcagtattgcaaatgaatttattgggacatgatttc
tttattttaccgatagagaaactgatggcacaagcattgtttatagacgaaaagatggc
aaatacggtctgattgaaacaactgaataa


Sequence 3000
MIRFEIHGDNLTITDAIRNYIEEKVGKLERYFNNVPNAVAHVRVKTYSNSTTKIEVTIPL
KDVTIRVLERHDDLYAGIDLITNKLERQVRKYKTRVNRKHRDRGDQDIFVAEVQESTTNN
HADDIESENDIEIIRSKQFSLKPMDSEEAVLQMNLLGHDFFIFTDRETDGTSIVYRRKDG
KYGLIETTE*

Sequence 3001
Contig_0562_pos_12885_13190
is similar to (with p-value 1.0e-20)
  >sp:sp|P46378|FAS6_RHOFA HYPOTHETICAL 21.1 KD PROTEIN IN FA
SCIATION LOCUS (ORF6). >pir:pir|F55578|F55578 hypothetical p
rotein 2 (ipt 3' region) - Rhodococcus fascians plasmid pFiD
188 >gp:gp|Z29635|RFCCIPTFD_6 R.fascians (D188) genes for P4
50 cytochrome, isopentenyltransferase and ferredoxine. NID:
g455000.
gtgttgaaaaagatacatgggttattgttttgtccaactcttgaccatgatacaacttcc
gttatttcttctaaagttccaggtccaccaggaagtgctatgcatacatcgcctttcttt
aagataacttctttttctttcagacatactttccacgattattagttcgtctaatttatca
tgtgctaactctctttgttttaaaaacgtaggcattacaccaattgttttttccgttgtta
tgtattacagtatttgcaattgttcccattaatcctgcatttccgccaccaaaaactaaa
gtgtga

Sequence 3002
VLKKIHGLLFCPTLDHDTTSVISSKVPGPPGSAMHTSPFFKITSFLSDILSTIISSSNLS
CANSLCFKNVGITPIVFPLLCITVFAIVPINPAFPPPKTKV*

Sequence 3003
Contig_0562_pos_13285_12725
is similar to (with p-value 8.0e-24)
>sp:sp|P46378|FAS6_RHOFA HYPOTHETICAL 21.1 KD PROTEIN IN FA
SCIATION LOCUS (ORF6). >pir:pir|F55578|F55578 hypothetical p
rotein 2 (ipt 3' region) - Rhodococcus fascians plasmid pFiD
188 >gp:gp|Z29635|RFCCIPTFD_6 R.fascians (D188) genes for P4
50 cytochrome, isopentenyltransferase and ferredoxine. NID:
g455000.
atgaatattatagtctattgtggagcaagcaagggaaacaaaaaagaatacgaaaatagt
gcaattcaattaggtgaatggatagctaaaaataatcacactttagtttttggtggcgga
aatgcaggattaatgggaacaattgcaaatactgtaatacataacaacggaaaaacaatt
ggtgtaatgcctacgtttttaaaacaaagagagttagcacatgataaattagacgaacta
ataatcgtggaaagtatgtctgaaagaaaagaagttatcttaaagaaaggcgatgtatgc
atagcacttcctggtggacctggaactttagaagaaataacggaagttgtatcatggtca
agagttggacaaaacaataacccatgtatctttttcaacacaaataattattattccctt
atcgaacaattctacgatcaaatggtttcaaacgagtttttaactcaagaagatagagat
aaaatattattctcaaactcattccaagaaattgaagaatttatagaaaactataaaaca
ccaaaaataagaacttattaa

Sequence 3004
MNIIVYCGASKGNKKEYENSAIQLGEWIAKNNHTLVFGGGNAGLMGTIANTVIHNNGKTI
GVMPTFLKQRELAHDKLDELIIVESMSERKEVILKKGDVCIALPGGPGTLEEITEVVSWS
RVGQNNNPCIFFNTNNYYSLIEQFYDQMVSNEFLTQEDRDKILFSNSFQEIEEFIENYKT
PKIRTY*

Sequence 3005
Contig_0563_pos_5497_5093
is similar to (with p-value 4.0e-71)
>gp:gp|AF046871|AF046871_3 Anabaena PCC7120 heterocyst-inhi
biting signaling peptide (patS) and holiday junction resolva
se (ruvC) genes, complete cds; and unknown genes. NID: g2896
023.
atgttagctaataatggtttaatcgcgattaatctcgcttatcagaatttagaaagagca
tttgttcaagatgtttctgatattgaatccaaacttacgttagcagcgacacctaagctc
gcatcaaaatcagctattagagaaagtatacgcttagcaattgttcctacaattgattct
gtaaaaacatatggtctagtttcaattccaggtatgatgacaggattgattatcggaggc
gttgacccacttcaagcaattaaatttcaattgcttgtcgtgtttattcatacaacagcg
acgattatgtctgcactcattgcaacgtatatgagttacggtcagttctttaatgctcgt
catcaactcattgctagaacgcaacgcacaagacaaagtagttaa

Sequence 3006
MLANNGLIAINLAYQNLERAFVQDVSDIESKLTLAATPKLASKSAIRESIRLAIVPTIDS
VKTYGLVSIPGMMTGLIIGGVDPLQAIKFQLLVVFIHTTATIMSALIATYMSYGQFFNAR
HQLIARTQRTRQSS*

Sequence 3007
Contig_0563_pos_1994_1008
is similar to (with p-value 2.0e-32)
>sp:sp|P77307|YBBM_ECOLI HYPOTHETICAL 28.2 KD PROTEIN IN US
HA-TESA INTERGENIC REGION.
atgaatgtagtacttattggtggtggcactggactttctgtccttgctagaggccttaga
gaatttccaatagacattactgccattgttactgtagcggacaatggtgggagcacgggg
aaaattagagatgtcatggatattccagcgcctggtgatattcgtaatgtcattgctgct
ttaagtgactcagaatcgatattaactcaattgttccagtaccgtttttggtgaaaatcaa
gtagatgggcattcattaggtaatttagttattgctggaatgactaacattactaatgat
tttggacacgctattaaagagttaagcaaagtttttaaatattaaaggccaagtcatccct
tcaacaaacgcaagtgtgcaactcaacgcggtgatggaagacggtgaaattgtacatgga

```
     gaaactaatatacctaaaacacataaaaaaatagatcgtgtgtttttagaaccaagtgat
     gttgaaccaatgaatgaagcgatagaagctttagaacaagcagatttaattgtcttagga
     ccaggttcattatatacaagtgttatatcaaatttatgtgtcaaaggtatttcagaagca
     ttattacgtacatctgctccaaaactttatgtatctaatgttatgacacaaccaggcgag
5    actgataattatgatgtcaaagagcatattgatgcacttactcgacaagttggtgaacca
     tttattgattttgtcatatgtagctcagaatcctatagtaaagatgtttttacaacgatat
     gaagaaaagaattcgaaaccagtagcagtacataaagaacaattaaaagatagtggaatt
     agagtttttaacggcatctaatttagttgaaatatctaatgaacactatgtcagacataac
     acaaaagtattatcaaaaatgatttatgaacttgccttagaattaacaagtacaattcgc
10   tttactcctagtgataaaaagaaataa
```

Sequence 3008
```
MNVVLIGGGTGLSVLARGLREFPIDITAIVTVADNGGSTGKIRDVMDIPAPGDIRNVIAA
LSDSESILTQLFQYRFGENQVDGHSLGNLVIAGMTNITNDFGHAIKELSKVLNIKGQVIP
STNASVQLNAVMEDGEIVHGETNIPKTHKKIDRVFLEPSDVEPMNEAIEALEQADLIVLG
PGSLYTSVISNLCVKGISEALLRTSAPKLYVSNVMTQPGETDNYDVKEHIDALTRQVGEP
FIDFVICSSESYSKDVLQRYEEKNSKPVAVHKEQLKDSGIRVLTASNLVEISNEHYVRHN
TKVLSKMIYELALELTSTIRFTPSDKKK*
```

Sequence 3009
Contig_0564_pos_5882_6601
is similar to (with p-value 7.0e-36)
   >sp:sp|P49309|MOCR_RHIME PROBABLE RHIZOPINE CATABOLISM REGU
LATORY PROTEIN MOCR. >pir:pir|S51574|S51574 mocR protein - R
hizobium meliloti >gp:gp|X78503|RMMOCCABR_6 R.meliloti mocC,
   ORF334, ORF293, mocA, mocB and mocR genes. NID: g468758.
```
     atgattatcgttgctacagctactgggggatatgcctttcccaagcgtggctaacatctta
     caagaaagactaggaacaggtaaggttgctactatggatcaacttgctgcatgttctggc
     tttatgtattcaatgattactgctaaacaatatatacaatctggtgattacaaacatatt
     ttagttgtgggtgctgataaaattatctaagattaccgatatgactgaccgttctactgct
     gtattatttggagacggtgctggagctgttgtcatgggagaagttgctgaaggtcgtggt
     atcattagctatgaaatgggttcagacggtagtggtggtaaatacttgtacttagataga
     gaaactggcaaactcaaaatgaatggtagagaagtatttaaatttgctgtgagaattatg
     ggtgatgcatctacgcgtgtagttgagaaagctggtttatcgtctgaagacatagactta
     tttgttccacatcaggctaatattagaattatggaatctgcgagagagagagattaggaata
     gaaagagagaaaaaatgagtgtctcagtaaataaatatggtaatacttcggctgcctcaata
     ccattaagtattaatcaagaattgcaaaatggaaaaatcaaagatgacgatactttagtc
     ttagttggcttcggtggaggtctaacttggggcgcaatcgttattaaatggggaaaatag
```

Sequence 3010
```
MIIVATATGDMPFPSVANILQERLGTGKVATMDQLAACSGFMYSMITAKQYIQSGDYKHI
LVVGADKLSKITDMTDRSTAVLFGDGAGAVVMGEVAEGRGIISYEMGSDGSGGKYLYLDR
ETGKLKMNGREVFKFAVRIMGDASTRVVEKAGLSSEDIDLFVPHQANIRIMESARERLGI
EREKMSVSVNKYGNTSAASIPLSINQELQNGKIKDDDTLVLVGFGGGLTWGAIVIKWGK*
```

Sequence 3011
Contig_0564_pos_6613_0
is similar to (with p-value 9.0e-47)
   >sp:sp|P30790|YHI3_RHOCA HYPOTHETICAL 33.7 KD PROTEIN IN HI
MA 5'REGION (ORF3). >pir:pir|C41608|C41608 hypothetical prot
ein 3 (himA 5' region) - Rhodobacter capsulatus >gp:gp|M8403
0|RCAHIMA_3 Rhodobacter capsulatus integration host factor (
himA) gene, complete cds. NID: g151940.
```
     atgaataaaaataatagagttgttataacgggtatcggagccttatctccaattggtaac
     gatgctaaaacaacatgggacaatgcactaaaaggtgttaacggtatagataaaatcaca
     agaatagatactgatgattataatgtacatcttgctggtgaattgaaagatttttaatata
     gaagaccacattgatagaaaagaagctcgccgtatggatcggtttacacaatacgcggtg
     gttgctgcaagagaagcggttaaagatgcacaattaaatattaatgaaaaaaatgccgac
```

```
cgtattggtgtatggattggttctggtatcggtggtatggaaactttcgaagttgcacat
acaacacttgtagaaagaggaccacgtcgagtaagtccatttttcgttccaatgttaatt
cctgatatggctactggtcaagtttctattgatttaggtgccaaagggcccaatggttct
acagtaacagcttgtgctacggggactaactcaataggtgaggcatttaaaattattcaa
cgtggtgatgcagatgcaatggtgactggtggaacagaggcacctattacacatatggca
atcgcagggtttagtgcaagtcgtgcattatctacaaacaatgaccctgaaacagcttgt
cgaccattccaagaaggccgtgatggctttgttatgggtgaaggggcaggtattgttgta
cttgaatcattagattcagctaaagagagaggcgctgaaatttacgctgaagttgtaggt
tatggttcctctggcgatgcacatcatattacagcacctgcgcctgaaggtgaaggtggc
tcacgagctatgcaagctgctttagatgatgctggaatcaaagctcaagatgtacagtat
ttaaatgcacatggcacaagtacacctgttggagatttatatgaggttcaagcgattaaa
aatacattcggtgatgctgcgaagtcattaaaagtaagttcaactaaatcaatgactgga
catttattaggtgctacaggtggaattgaagctattttttctgcgctatcaattcgtgat
tcaaaggtagcccctacaatacatgcaatcacaccagacgaagaatgtgatttg
```

Sequence 3012

MNKNNRVVITGIGALSPIGNDAKTTWDNALKGVNGIDKITRIDTDDYNVHLAGELKDFNI
EDHIDRKEARRMDRFTQYAVVAAREAVKDAQLNINEKNADRIGVWIGSGIGGMETFEVAH
TTLVERGPRRVSPFFVPMLIPDMATGQVSIDLGAKGPNGSTVTACATGTNSIGEAFKIIQ
RGDADAMVTGGTEAPITHMAIAGFSASRALSTNNDPETACRPFQEGRDGFVMGEGAGIVV
LESLDSAKERGAEIYAEVVGYGSSGDAHHITAPAPEGEGGSRAMQAALDDAGIKAQDVQY
LNAHGTSTPVGDLYEVQAIKNTFGDAAKSLKVSSTKSMTGHLLGATGGIEAIFSALSIRD
SKVAPTIHAITPDEECDL

Sequence 3013
Contig_0564_pos_1547_414
  >pir:pir|I41060|I41060 3-oxoacyl-[acyl-carrier-protein] syn
thase (EC 2.3.1.41) - Escherichia coli >pir:pir|I84544|I8454
4 beta-ketoacyl-acyl carrier protein synthase II - Escherich
ia coli >gp:gp|U20767|ECU20767_1 Escherichia coli beta-ketoa
cyl-acyl carrier protein synthase II (fabF) gene, complete c
ds. NID: g664869. >gp:gp|AE000210|AE000210_5 Escherichia col
i K-12 MG1655 section 100 of 400 of the complete genome. NID
: g1787332. >gp:gp|Z34979|ECFABJ_1 E.coli fabJ gene encoding
 beta ketoacyl-acyl carrier protein synthase. NID: g510831.
gtgatacatcgacaaacaaatcaattaaacactgataacttagaaaagaagcaacgacaa
tataaatatgcttttaatttagctgaaattgattctgaaagtttttcctatgcacatttt
agaaaatatgcgaaagacgttttttgaagaccatcagctatcactattacaacgaggcgag
cgtcaaggggaatatattttaaggcaacaaatttcacattacttatttaatagtcgtggc
gtcacttgtcacccaaatcaaattattgttggatcatcaacaagccagttactcgatatg
ataaccaatttactaaaaaaagaagaatttattattgaacagccaagttatccacctatt
aaacatacgcttgataaaaaaggtataagttatattcaagtcccagttgaacaaaatgga
atacaaatcgaccctattttaaatacaaataacaatattttgtatataacaccatctcat
caatttccaactggttatgtcaccaatttaaaaaaaagaacacaattaatcaattggtcc
catcaagctaagcaaagatatattatcgaagatgattatgattcagaatttagatatttt
ggcaaacccatacctgcattacaaagtttagacacaaaaggaaaagtcatttatattagt
actttctcaaaatctttatatccaagctgtaggattgcatatattgttttgccacaaaat
ttaatgcacaaatataataatcaaaaatataaagagggaaatacagtgcctgtgcatgtt
caacacatgattgctcaattcatgataagtgggaaatttgaaagacatttgaataaaatg
cgaaagatatatagagataaacttgattatattttaaaacgattaaagccctacaatact
caaaaaagattgaaggcgcactaactggaatgcattttacaataactgttaatatgga
ttgtcaatgaaacaatgtttaaaaaatgcgaaaaaaaataatttaaaattaaaaccttat
cattacgaaaattattctaaagtttatccaaaatttattttaggatttggggggataaaa
aaagaagaattagaagatcatgttaatgcattaattcattcactcgttatataa
```

Sequence 3014
VIHRQTNQLNTDNLEKKQRQYKYAFNLAEIDSESFPMHIFRKYAKDVFEDHQLSLLQRGE
RQGEYILRQQISHYLFNSRGVTCHPNQIIVGSSTSQLLDMITNLLKKEEFIIEQPSYPPI
KHTLDKKGISYIQVPVEQNGIQIDPILNTNNNILYITPSHQFPTGYVTNLKKRTQLINWS

HQAKQRYIIEDDYDSEFRYFGKPIPALQSLDTKGKVIYISTFSKSLYPSCRIAYIVLPQN
LMHKYNNQKYKEGNTVPVHVQHMIAQFMISGKFERHLNKMRKIYRDKLDYILKRLKPYNT
QIKIEGALTGMHFTITVNNGLSMKQCLKNAKKNNLKLKPYHYENYSKVYPKFILGFGGIK
KEELEDHVNALIHSLVI*

Sequence 3015
Contig_0566_pos_9236_0
is similar to (with p-value 5.0e-35)
 >sp:sp|P94378|3MGH_BACSU PUTATIVE 3-METHYLADENINE DNA GLYCO
SYLASE (EC 3.2.2.-). >gp:gp|Z99123|BSUB0020_157 Bacillus sub
tilis complete genome (section 20 of 21): from 3798401 to 40
10550. NID: g2636240. >gp:gp|D83026|D83026_59 Bacillus subti
lis genome sequence covering lic-cel region. NID: g1783231.
atgaatatagttgatgtaaatgatgttgagaaaatcgacttagcaatagatggcgctgat
gaagtagacagtgcgcttaaccttattaaaggcggtggtggagccttatttagggaaaag
gtcatagatgaaatggctgaccgatttgtcgttgttgtagatgaaagtaaactcgtcaac
tatttaggagaaacatttgcattaccagttgaagtcgataaatttaattggtaccaagtt
gccaaaaaaattgagcgtacttatgatattcatgtaagcagaagagttaatgaagatgta
ccgtttataaccgacaatggtaattacatattagattgttcattgcaaaatagaattcct
gcttatgagctac

Sequence 3016
MNIVRVNDVEKIDLAIDGADEVDSALNLIKGGGGALFREKVIDEMADRFVVVVDESKLVN
YLGETFALPVEVDKFNWYQVAKKIERTYDIHVSRRVNEDVPFITDNGNYILDCSLQNRIP
AYELX

Sequence 3017
Contig_0566_pos_8875_8186
is similar to (with p-value 3.0e-19)
 >sp:sp|P32157|YIIM_ECOLI HYPOTHETICAL 26.6 KD PROTEIN IN KD
GT-CPXA INTERGENIC REGION (O234). >pir:pir|S40854|S40854 hyp
othetical protein o234 - Escherichia coli >gp:gp|L19201|ECOU
W87_53 E. coli chromosomal region from 87.2 to 89.2 minutes.
 NID: g304961. >gp:gp|AE000466|AE000466_1 Escherichia coli K
-12 MG1655 section 356 of 400 of the complete genome. NID: g
2367328.
gtgatatacatgattaaagtgaatgccatatctattggcaaaatagaaacattgtcttat
ggaaactataaaccaatgcaatcagcgttaaacaaaattccttttaaaggtcaaatgtgg
ctcaatcgacttgggttcgtggacgatgaacaagcctatcataaccatggtggtatacat
aaagcgatatgttgttttagtaaatctaattatcaattatttaaagatgacttagatcaa
ttacctgaatttgcaatgtttggagagaatttgacagttgaacatctagatgaagcagat
gtttatttttggtaatcagtatcaactaggcgatacaatcatagaagtatcagatatacga
gaaccttgttggaaaattcaagctaaatatgcaatacctaatttagttcaaaaaagtcg
caatctggtaaaactggattttattttagagttataaaagaaggatatgtacatcagagt
gataatttaaaactcattaaaaaggcagaatcaaacacacgtctatctgtgaaagactta
aatcatctattctataatgagcgaaataatttaagattaatctatcatgcacttcgaaat
ccttatctttcacctgatcgaaagaaaaaactacagaaaatgaaacgcgtgccgaaaat
agaaaattcattaaatctgacgataaatag

Sequence 3018
VIYMIKVNAISIGKIETLSYGNYKPMQSALNKIPFKGQMWLNRLGFVDDEQAYHNHGGIH
KAICCFSKSNYQLFKDDLDQLPEFAMFGENLTVEHLDEADVYFGNQYQLGDTIIEVSDIR
EPCWKIQAKYAIPNLVQKMSQSGKTGFYFRVIKEGYVHQSDNLKLIKKAESNTRLSVKDL
NHLFYNERNNLRLIYHALRNPYLSPDRKKKLQKMKTRAENRKFIKSDDK*

Sequence 3019
Contig_0566_pos_3563_2946
is similar to (with p-value 8.0e-19)
 >sp:sp|P47968|RPIA_MOUSE RIBOSE 5-PHOSPHATE ISOMERASE (EC 5

.3.1.6) (PHOSPHORIBOISOMERASE)..>gp:gp|L35034|MUSRPI_1 Mus m
usculus ribose 5-phosphate isomerase (RPI) mRNA exons 1-6, c
omplete cds. NID: g836673.

```
gtgatatatttggactttataaaatcagcagacaacccaaactgcaaaagctttattaggt
5   gttaaaattatttaccaagatgactatcaaacatatactggatatattgtagaaactgaa
gcttatttaggtatacaagataaagctgcacatggttttggtggcaaaataacaccaaaa
gtgacttctttatatataaaaaaggtggcacgatatatgcacatgtgatgcatacgcactta
ttaatcaattttgttacacggactgagggcataccagaaggtgtacttattcgtgctatt
gaaccagatgaaggtatcggcgctatgaacgtcaatcgtggaaaatctggatacgagctc
10  actaatggtccaggaaagtggactaaagctttcaatattccacgatcaattgatggctca
accttaaatgactgcaaattatctatagataccaatcatcgcaaatatccaaaaactatt
atagaaagtggtcgtatcggtattcctaataaaggagaatggacaaataaaccactacgt
ttcactgttaaaggcaatccatatgtctctagaatgcgcaaatcagattttcaaaatccc
gacgatacatggaaataa
```

15

Sequence 3020
VIYLDFINQQTTQTAKALLGVKIIYQDDYQTYTGYIVETEAYLGIQDKAAHGFGGKITPK
VTSLYKKGGTIYAHVMHTHLLINFVTRTEGIPEGVLIRAIEPDEGIGAMNVNRGKSGYEL
TNGPGKWTKAFNIPRSIDGSTLNDCKLSIDTNHRKYPKTIIESGRIGIPNKGEWTNKPLR
20 FTVKGNPYVSRMRKSDFQNPDDTWK*

Sequence 3021
Contig_0568_pos_4867_2714
is similar to (with p-value 3.0e-54)
25 >sp:sp|P45112|RECJ_HAEIN SINGLE-STRANDED-DNA-SPECIFIC EXONU
CLEASE RECJ (EC 3.1.-.-). >pir:pir|F64110|F64110 single-stra
nded-DNA-specific exonuclease (recJ) homolog - Haemophilus i
nfluenzae (strain Rd KW20) >gp:gp|U32801|U32801_1 Haemophilu
s influenzae Rd section 116 of 163 of the complete genome. N
30 ID: g1574143.

```
gtggataatcaggagattcaaaatctatttgaaggcactaacattagtcatgattatatg
ttattaagtgatatgcaaaaagccattgatcgtattaaaattagctatcgatcaaaatgaa
cgaatactagtatatggtgactatgatgcagatggtgttacatctactacaattctagtg
tctactttacgtctacttggcgctcaggtggggtggtatattcccaatagatttacagaa
35  ggatatggacctaatgaattagcatttaaaaatgcttatgacgaagggatttccttaata
ataactgtagataatggtatacagggacatcatgaaataagtaccatacaagaattaggt
gtagatgttatagtgacagatcatcatgaaataggagaaactttacctgatgcttttgca
attgtacatccgatgcatcctaattttgaatatccttttaaatatttatgtggtgcgggt
gttgcttataaattggcccaaggattgatagagcatccacctcaacatttcatagcatta
40  gctgccataggtacaattgcagatttagtatcattgacagatgagaatagatatattgta
aagcaaggagtaaagatattaaacaatcatcacaccatcgtccataaaggctatcttaaat
caagcggggtttttaatgatgaaataacagaagaaacaattggttttattattggacctcga
ttaaatgcggttggtagactagaagatgcatcattagctgctgaactttttattgtccgat
gaatttgaagaggcggaatttttagctgaacaagttgaacattttaatcatgaacgtaaa
45  gatatagtatctaaaattactaatgaagcattgttattagcagaggaacaaatcaagcaa
ggccatttgtttctttttacttgccaaagaggggtggcatgagggtgtattaggtattgtt
gcatctaaaattgtagaaacatatgcactacctacattaatttaatatgatgaaaat
caaaatcatgccaagggttctgcgaggtcgattgaacaagttccatgtttgatattta
aatgatcatcaacagttaattgataagtttggtggtcatcacatggctgcaggaatgaca
50  atgtctatcgataatattgaacatttacataaagagctagatatgtggatgaaagaacta
actgttaccacttcattagagccttcaataaaggtggatgcacaacttgaagaaaaagaa
attaacattaaaaatattaaagatatatttcaattaaggcctttggtacggactttaat
agtcctctttttatggttagagatctaattgtcaagtcaacaaagggaattggacaggat
aataagcatcttaagttaacgcttggtcattcaggtttaactgctttattttggaatcat
55  ggacatttggcaagtgaacttgaaccaggtcaaccgattcatataataggaacattgcaa
attaatgaatggaatggtaatcaaacacctcaatttatcatcaaagacattgctatagac
caattacaaattttagattatcgtagtaaacgtaaaaatatacaatttaaagaaactgaa
tcaaatgttgcctatgtcattcatccgaaacttaaaaaaagcaattcacattactatcat
tatggtgaggttattgatagaccttatgataaaatagtgtttagagatttacctaatact
```

```
atggttgaaattgaacaaaccttagaacattcacaaatttctcaactatatttagttctg
cagcatgaaaagtcaatatattttgaaggtatccctagcaagtcgctttttaaaaagtgt
tacaaagctttaataaacaaaaaagaaactgacttaattaaagaaggtatgctgctttgt
gagtacttaaatattaagccagaaatattaacattcatgcttaaagtctttaaagaattg
gagtttatttacgacgaaaagggattgattaagataaatccagcgccaaataaacaagat
atagaaaatagtcgtatttaccaaatgagacgagcacgtatggaagtagaagaacgtctt
ctctatgatgatttttttaaatataaaagaatggataatatcaaagttaacatag
```

Sequence 3022
```
VDNQEIQNLFEGTNISHDYMLLSDMQKAIDRIKLAIDQNERILVYGDYDADGVTSTTILV
STLRLLGAQVGWYIPNRFTEGYGPNELAFKNAYDEGISLIITVDNGIQGHHEISTIQELG
VDVIVTDHHEIGETLPDAFAIVHPMHPNFEYPFKYLCGAGVAYKLAQGLIEHPPQHFIAL
AAIGTIADLVSLTDENRYIVKQGVKILNNHTPSSIKAILNQAGFNDEITEETIGFIIGPR
LNAVGRLEDASLAAELLLSDEFEEAEFLAEQVEHFNHERKDIVSKITNEALLLAEEQIKQ
GHLFLLLAKEGWHEGVLGIVASKIVETYALPTLILNIDENQNHAKGSARSIEQVSMFDIL
NDHQQLIDKFGGHHMAAGMTMSIDNIEHLHKELDMWMKELTVTTSLEPSIKVDAQLEEKE
INIKNIKDIFQLRPFGTDFNSPLFMVRDLIVKSTKGIGQDNKHLKLTLGHSGLTALFWNH
GHLASELEPGQPIHIIGTLQINEWNGNQTPQFIIKDIAIDQLQILDYRSKRKNIQFKETE
SNVAYVIHPKLKKSNSHYYHYGEVIDRPYDKIVFRDLPNTMVEIEQTLEHSQISQLYLVL
QHEKSIYFEGIPSKSLFKKCYKALINKKETDLIKEGMLLCEYLNIKPEILTFMLKVFKEL
EFIYDEKGLIKINPAPNKQDIENSRIYQMRRARMEVEERLLYDDFLNIKEWIISKLT*
```

Sequence 3023
Contig_0569_pos_4529_5401
is similar to (with p-value 3.0e-33)
    >sp:sp|P39312|CYCA_ECOLI D-SERINE/D-ALANINE/GLYCINE TRANSPO
    RTER. >pir:pir|S56433|S56433 hypothetical protein o470 - Esc
    herichia coli >gp:gp|U14003|ECOUW93_120 Escherichia coli K-1
    2 chromosomal region from 92.8 to 00.1 minutes. NID: g126317
    2. >gp:gp|AE000492|AE000492_4 Escherichia coli K-12 MG1655 s
    ection 382 of 400 of the complete genome. NID: g1790649.
```
atgcctgaactacctgaagttgaacatgttaaaagaggtattgagccatttataaaaagt
gcaaaaatagagaaagtaacttttgctaaaaatgtaattaacggtaagaataataaccgt
gagacgattataaaaggtatggaattagatacttttaaaaaacttactgaaggttatgtt
ataaaaaaagttgaaagaagaagtaagtacattattttttatatagcggatcatgacgat
gatagaatcttagttagtcatttaggtatggcaggcggattctttgttgttaataacctt
gatgagataagtacaccgaattatcgaaagcattggcaagtcattttcgatttggataat
aaacaaaaattagtctattctgatatcagacggtttggagaaattagaaatatagtcaat
tttgatagttatccatctttattagaaatcgctccagaaccatttgaagaggtggcattt
gaacactatttagaatgtttgacaatgaaaaaatataagaataaaccaataaacaaacg
attcttgatcatcgtgttatagcaggagctggaaatatctatgcctgtgaagctttattc
agagctggtattactccggataaaattactaattcactcactaaacaagaaagaaaatcc
ctctttattatgttcgagaagtttttagaagagggtataaaatatggaggtactagtatt
tcagattataggcatgcagatggtaaaactggacaaatgcaattacatttaaatgtatat
aaacaaaaaaagtgcaaggtttgtggtcattcgattgaaacaaaagtgatagctggtaga
aatagtcatttttgcccaaactgtcagagataa
```

Sequence 3024
```
MPELPEVEHVKRGIEPFIKSAKIEKVTFAKNVINGKNNNRETIIKGMELDTFKKLTEGYV
IKKVERRSKYIIFYIADHDDDRILVSHLGMAGGFFVVNNLDEISTPNYRKHWQVIFDLDN
KQKLVYSDIRRFGEIRNIVNFDSYPSLLEIAPEPFEEVAFEHYLECLTMKKYKNKPIKQT
ILDHRVIAGAGNIYACEALFRAGITPDKITNSLTKQERKSLFYYVREVLEEGIKYGGTSI
SDYRHADGKTGQMQLHLNVYKQKKCKVCGHSIETKVIAGRNSHFCPNCQR*
```

Sequence 3025
Contig_0569_pos_8390_8887
is similar to (with p-value 1.0e-38)
    >sp:sp|P44948|FPG_HAEIN FORMAMIDOPYRIMIDINE-DNA GLYCOSYLASE
    (EC 3.2.2.23) (FAPY-DNA GLYCOSYLASE). >pir:pir|A64104|A6410

4 formamidopyrimidine-DNA glycosylase (fpg) homolog - Haemop
hilus influenzae (strain Rd KW20) >gp:gp|U32776|U32776_1 Hae
mophilus influenzae Rd section 91 of 163 of the complete gen
ome. NID: g1573969.

5   gtgattaaacttgggagggtgacatatatgaaatgcccaaaatgtaattctacacattcc
        agagtggttgattcaagacatgcagatgaggccaatgcgattagacgtagaagagaatgt
        gaaaattgcggaacgcgtttacaacatttgaacatattgaagttagtccattaatagta
        gtgaagaaagatgggactagagaacaatttttaagagaaaaaatattaaatggtttagta
        agatcttgcgagaaacgaccagtacgttatcaacaacttgaagacataactaataaagtg
10   gagtggcaacttagagatgagggacaaactgaaatttcttctagagaaattggagagcat
        gttatgaatttgttaatgcatgttgaccaagtttcctatgtaagatttgcatctgtatat
        aaagaattcaaagatgttgatcaactcttagagtcaatgcaaggtatcttgagtgataat
        aaacggagtgataaatag

15   Sequence 3026
        VIKLGRVTYMKCPKCNSTHSRVVDSRHADEANAIRRRRECENCGTRFTTFEHIEVSPLIV
        VKKDGTREQFLREKILNGLVRSCEKRPVRYQQLEDITNKVEWQLRDEGQTEISSREIGEH
        VMNLLMHVDQVSYVRFASVYKEFKDVDQLLESMQGILSDNKRSDK*

20   Sequence 3027
        Contig_0569_pos_1314_691
        is similar to (with p-value 1.0e-33)
         >sp:sp|P46227|YRS1_SYNP6 HYPOTHETICAL 19.1 KD PROTEIN IN PS
        BB-RPS1 INTERGENIC REGION (ORF 168). >pir:pir|S51484|S51484
25   hypothetical protein 168 - Synechococcus sp. >gp:gp|D28752|S
        YORPS1_2 Synechococcus sp. gene for ribosomal protein S1, co
        mplete cds. NID: g560122.
        atgtctgttgtaccatggcaacaattgaatcctgctgacagtccatacgttaaaatgttt
        ggattagttggaatcccttttgcagcaggtattattaactttgttgtacttacagctgca
30   gcctcttcttgtaatagtggtatatttgctaatagccgtacgatgtttggattagctgga
        agaaagcaaggtccagcattcttacatagaaccaataagcacggcgtaccacattatgct
        attttagtgacatgtggcttattaagtatttcagtcgtgttaaatgcaatttttaaagat
        gcgactaaagtgttcgtacaaattacaacatttcaactgtttaaatattatgatttgg
        acaattattatgatcgcgtatctaggttatttaagacatgaaccgaaacagcataaagaa
35   agtaactataaaatgtggggcggaaaatacatggcttacagtattttagggttctttgca
        tttattttattatactattgattaatagtgcaacgcgttatgccgtactttctgcaccc
        gtatggtttgttatcatgctattgatgtatcaaaaatataaaaaagaatctcgcaaagct
        aaaattaaaaatgaggaagagtaa

40   Sequence 3028
        MSVVPWQQLNPADSPYVKMFGLVGIPFAAGIINFVVLTAAASSCNSGIFANSRTMFGLAG
        RKQGPAFLHRTNKHGVPHYAILVTCGLLSISVVLNAIFKDATKVFVQITTFSTVLNIMIW
        TIIMIAYLGYLRHEPKQHKESNYKMWGGKYMAYSILGFFAFIFIILLINSATRYAVLSAP
        VWFVIMLLMYQKYKKESRKAKIKNEEE*
45
        Sequence 3029
        Contig_0570_pos_0_488
        is similar to (with p-value 5.0e-39)
         >sp:sp|P54715|PTIB_BACSU PTS SYSTEM, ARBUTIN-LIKE IIBC COMP
50   ONENT (PHOSPHOTRANSFERASE ENZYME II, BC COMPONENT) (EC 2.7.1
        .69). >gp:gp|Z99108|BSUB0005_89 Bacillus subtilis complete g
        enome (section 5 of 21): from 802821 to 1011250. NID: g26330
        55. >gp:gp|D50543|D50543_3 Bacillus subtilis DNA for 76-degr
        ee region, complete cds. NID: g1486240.
55   atgtttgctttttcggtattgtttgggattcgctacattattttaaaaatccaaccatt
        atgggaggattagctgatcagcaaacattttggtttaaattttggtctgttattgaatca
        ggtggttgggtaatatttacacatatggaaattgtctttgtagttggcttaccattatct
        cttgctaaaaaagcaccaggacatgcagctttagcagctctaatgggatatttaatgttt
        aatacttttatcaatgcaattttaactcaatggccacatactttggcgctaatttaaaa

753

aaaggtgtagaaaacacaacaggattaaaatcgattgcaggtattgaaacgttagatacc
aatattttaggtgcaatcattatctcaggaataataacgtggatacataatagatattac
agtaagcgtttacctgaaatgttaggtgtatttcaaggattaacattcgttgtaacaatc
tctttctt

Sequence 3030
MFAFFGIVLGFATLFKNPTIMGGLADQQTFWFKFWSVIESGGWVIFTHMEIVFVVGLPLS
LAKKAPGHAALAALMGYLMFNTFINAILTQWPHTFGANLKKGVENTTGLKSIAGIETLDT
NILGAIIISGIITWIHNRYYSKRLPEMLGVFQGLTFVVTISFX

Sequence 3031
Contig_0580_pos_966_1382
is similar to (with p-value 5.0e-85)
  >sp:sp|P43848|PUR5_HAEIN PHOSPHORIBOSYLFORMYLGLYCINAMIDINE
CYCLO-LIGASE (EC 6.3.3.1) (AIRS) (PHOSPHORIBOSYL-AMINOIMIDAZ
OLE SYNTHETASE) (AIR SYNTHASE). >pir:pir|G64122|G64122 5'-ph
osphoribosyl-5-aminoimidazole synthetase (purM) homolog - Ha
emophilus influenzae (strain Rd KW20) >gp:gp|U32822|U32822_2
 Haemophilus influenzae Rd section 137 of 163 of the complet
e genome. NID: g1574265.
atggaagcacaaatggaaaaagatggtaattactatatggaaggaatattagatgatatt
caacaagatggatatggtttcttaagaaccgttaactattctaaaggtgagaaggatatt
tatatttctgcaagccaaattcgacgttttgaaataaaacgtggtgataaagtaacgggt
aaagttcgtaaaccaaaagataatgaaaaatattatggtctacttcaagttgattttgta
aacgaccataatgcagaagaagtcaaaaaacgtccttcacttccaagctttaacacctct
ttatccggaagaaagaatcctattagaaacgcaatctacaaattattccactcgtattat
ggatttagtcacaccaataggtcttggtcaacgtggtcttatagttgcaccacctaa

Sequence 3032
MEAQMEKDGNYYMEGILDDIQQDGYGFLRTVNYSKGEKDIYISASQIRRFEIKRGDKVTG
KVRKPKDNEKYYGLLQVDFVNDHNAEEVKKRPSLPSFNTSLSGRKNPIRNAIYKLFHSYY
GFSHTNRSWSTWSYSCTT*

Sequence 3033
Contig_0580_pos_2122_3000
is similar to (with p-value 1.0e-27)
  >sp:sp|P38527|RHO_THEMA TRANSCRIPTION TERMINATION FACTOR RH
O. >gp:gp|L27279|TMORHO_1 Thermotoga maritima rho gene, comp
lete cds. NID: g454858.
atgaaagcgccagttctggtatcaggtactgacggtgtgggtacaaagttaaaattagca
attgactatggaaagcatgacacaattggtattgatgctgtcgcaatgtgtgtaaatgat
attttaacaacaggtgctgaacctttatacttttttagactatattgccacgaataaagta
gtgccaagtactatagagcaaatcgttaaaggtataagtgacggttgcgaacaaaccaat
acggcacttataggcggtgaaactgctgaaatgggagaaatgtatcatgaaggtgaatat
gatattgctggtttttgcagtaggagcggtagagaaagaggactatattgatggttcaaat
gttgaagaaggacaagcaattattggtttagcttcaagtggtattcattcaaatggctat
agtctagttagaaaaatgataaaagaatcaggagttcaattacatgatcaatttaatggt
caaacctttttagaaaccttccttgcaccaacaaaattgtatgtaaagcctattcttgaa
ttaaagaaacatattgatatcaaagcgatgagccatattactggtggaggtttctatgaa
aatattccgcgtgcccttcctaaaggtttatcagcaaaaatagatacacaatcattccca
acgttggaagtctttaattggcttcaaaaacagggcaacatttcaacgaatgaaatgtat
aacatatttaatatgggtattggatatacaattattgttgacaaaaaagatgttcaaaca
acattaacaacgttacgtgcaatggatacaactgcatatgaaattggtgagattataaaa
gatgatgatacacctattcatttattggaggtagaatag

Sequence 3034
MKAPVLVSGTDGVGTKLKLAIDYGKHDTIGIDAVAMCVNDILTTGAEPLYFLDYIATNKV
VPSTIEQIVKGISDGCEQTNTALIGGETAEMGEMYHEGEYDIAGFAVGAVEKEDYIDGSN
VEEGQAIIGLASSGIHSNGYSLVRKMIKESGVQLHDQFNGQTFLETFLAPTKLYVKPILE

LKKHIDIKAMSHITGGGFYENIPRALPKGLSAKIDTQSFPTLEVFNWLQKQGNISTNEMY
NIFNMGIGYTIIVDKKDVQTTLTTLRAMDTTAYEIGEIIKDDDTPIHLLEVE*

Sequence 3035
Contig_0581_pos_6447_3826
is similar to (with p-value 4.0e-19)
  >gp:gp|X99710|LLLVSFPEP_4 L.lactis ORF, genes homologous to
  vsf-1 and pepF2 and gene encoding protein homologous to met
  hyltransferase. NID: g1771200.
atgaaagctagtgaaattaggcaaaaatatttgaatttctttgtagaaaaaggacatatg
attgaaccgtctgcaccacttgtacctatcgatgatgattcattattgtggatcaattct
ggtgtagctaccttgaaaaaatactttgacggacgcgaaactcctaaaaaaccaagaatt
gtcaattctcaaaaagctatacgaacaaacgatatagaaacgtcggtttactgctcgc
catcatacttttttgagatgctaggtaattttcaatcggtgactactttaaacatgaa
gcgattgaatttgcttgggaatttctaacaagtgataaatggatgggcatggaacctgag
aagctatatgtgactattcatcctgaagatactgaagctttagaatttggcatgaagac
attggtttagaggaaagtcgcatcattcgaatagaaggcaatttttgggatatcggggaa
ggaccatctggaccaaatacagaaattttctatgatcgtggatcggcttatggaaaagat
gatcctgctgaagaaatgtatcctggtggtgaaaatgaaagatatttagaagtttggaat
ctagtgtttagtgagtttaatcacaataaagacaatacttacacaccactaccaaataaa
aatattgatactggcatgggattagaacgtatgacgtctatctcacaaaatgtaagaaca
aattatgaaacagacttatttatgcctataattaaggaagtagaacatgtttcaggaaaa
aaatatttaattgatgatgcacaagatgttgcatttaaagttatcgcagaccacattaga
acgatttctttcgcaattgctgatggcgcattaccagctaatgagggtagaggatatgta
ttaagaagattattacgcagagcagttcgctttagccaatcattaggaattaatgaacca
tttatgtataaacttgttgatatagtcgctgatatcatggaaccatattatccaaatgtc
aaagacaaatccaactttattaaacgtgtcattaaatcagaggaagaacgcttccatgaa
acgcttgaggaaggtcttacgatttttaaatgaactgataaaagaagcgaagaatagtgat
caggttattaaaggtcatgatgcttttaagttatatgatacttatggattcccaatagaa
ttaactgaagaattagcaactcaagaaaaatttgtctgttgatatgcctacttttgaacag
gaaatgcaacaacagagagatcgagctagacaagctcgacagaattctcaatcaatgcaa
gttcaaagtgaagtactaaaaaatattcaagatgaaagtcaatttgttggctatgaaact
acggactatcaatcattaataactcatatcatatacaatggtgaagaagttaaacatgtt
gaagcaggagaaacaatttactttattttaagagaaacgcctttctatgcagtaagtggt
ggacaggtcgcagataagggaacagttggtaatgagagctttgaaataaatgtaactgac
gtaactaaagcgcctaatggccaaaacttacacaaaggtattgtgcaatttggtgaagca
acacagaacgcgaaagtagaagcacgtgttaacaaagaggatagacgacttattcaaaaa
aatcatagtgctacacatttattacatgctgcattaaaagaagtattaggagatcatgtt
aatcaggctggttcgttagtagaacctgaaagactacgttttgatttctcacattttggt
cctatgacacaagaagaaattaatttagtagaacgtagagtaaatgaagaaatttggaga
gctatcgacgtccgtattcaagaaatgagtattgaagaagccaaatcaataggcgctatg
gctttatttggtgaaaaatatggagatattgttcgcgttgtcaatatggcaccattttca
atagaattgtgcggtggaatacatgtgaataatactgcggaaattggtctctttaagatt
gtgagtgaatctggaacaggtgccggtgttagaagaattgaagctttaacaggtaaggt
gcattcttacatcttgaagaaattgaaacacagtttaataatattaaaaatcatttaaaa
gttaaatccgataaccaagtagttgaaaaagttaaacaacttcaagaagaagaaaaagga
ctgcttaaacaattagaacaacgcaacaaagaataacatcactaaagatggggaacatt
gaagaacaggttgagttgattaataatttgaaagtttttagcaacagaagtagaaattcca
aatccaaaagcgatacgttcaactatggatgactttaaatctaaacttcaagatactatt
atagtgttagtcggacaagttgatggaaaggtttctgtaattgctacagtaccaaaatca
cttacaaatcaagtaaaagctggagatcttatcaaaaacatgacaccaattattggtgga
aaaggtggaggtcgtcctgatatggctcaaggtggcggaactcaacctgaaaaaataaca
gaagcattacgctttattaaagattacattaaaaatctataa

Sequence 3036
MKASEIRQKYLNFFVEKGHMIEPSAPLVPIDDDSLLWINSGVATLKKYFDGRETPKKPRI
VNSQKAIRTNDIENVGFTARHHTFFEMLGNFSIGDYFKHEAIEFAWEFLTSDKWMGMEPE
KLYVTIHPEDTEAFRIWHEDIGLEESRIIRIEGNFWDIGEGPSGPNTEIFYDRGSAYGKD
DPAEEMYPGGENERYLEVWNLVFSEFNHNKDNTYTPLPNKNIDTGMGLERMTSISQNVRT

NYETDLFMPIIKEVEHVSGKKYLIDDAQDVAFKVIADHIRTISFAIADGALPANEGRGYV
LRRLLRRAVRFSQSLGINEPFMYKLVDIVADIMEPYYPNVKDKSNFIKRVIKSEEERFHE
TLEEGLTILNELIKEAKNSDQVIKGHDAFKLYDTYGFPIELTEELATQENLSVDMPTFEQ
EMQQQRDRARQARQNSQSMQVQSEVLKNIQDESQFVGYETTDYQSLITHIIYNGEEVKHV
EAGETIYFILRETPFYAVSGGQVADKGTVGNESFEINVTDVTKAPNGQNLHKGIVQFGEA
TQNAKVEARVNKEDRRLIQKNHSATHLLHAALKEVLGDHVNQAGSLVEPERLRFDFSHFG
PMTQEEINLVERRVNEEIWRAIDVRIQEMSIEEAKSIGAMALFGEKYGDIVRVVNMAPFS
IELCGGIHVNNTAEIGLFKIVSESGTGAGVRRIEALTGKGAFLHLEEIETQFNNIKNHLK
VKSDNQVVEKVKQLQEEEKGLLKQLEQRNKEITSLKMGNIEEQVELINNLKVLATEVEIP
NPKAIRSTMDDFKSKLQDTIIVLVGQVDGKVSVIATVPKSLTNQVKAGDLIKNMTPIIGG
KGGGRPDMAQGGGTQPEKITEALRFIKDYIKNL*

Sequence 3037
Contig_0581_pos_3500_3072
is similar to (with p-value 5.0e-25)
   >gp:gp|U46071|RCU46071_3 Rhodobacter capsulatus cytochrome
c biogenesis (cycH) gene, complete cds, and sarcosine oxidas
e gene, partial cds. NID: g1353871.
   atgctaaagcataaaattttaggactagatgttggaagtaaaactgttggcattgctata
   agtgaccttatgggttggactgctcaagggttagacacactccgtatcaacgaagaacaa
   gatgatttaggaattgatcaactcgtgaagattattaaagataatcaagttggaactgtt
   gttattggcttgcccaagaatatgaacaattcaataggtttttcggggagaggcttcaata
   aaatataaagaaaagttacaagagtctatcccttctattgatattgttatgtgggacgaa
   cgtttaagtacaatggctgctgaaagatctttacttgaagcagatgtttcaagacaaaaa
   agaaaaaaggtaatagataaaatggcagctgtatttattttacaaggctatttagattct
   attcaataa

Sequence 3038
MLKHKILGLDVGSKTVGIAISDLMGWTAQGLDTLRINEEQDDLGIDQLVKIIKDNQVGTV
VIGLPKNMNNSIGFRGEASIKYKEKLQESIPSIDIVMWDERLSTMAAERSLLEADVSRQK
RKKVIDKMAAVFILQGYLDSIQ*

Sequence 3039
Contig_0581_pos_2179_1685
   >sp:sp|P00957|SYA_ECOLI ALANYL-TRNA SYNTHETASE (EC 6.1.1.7)
   (ALANINE--TRNA LIGASE) (ALARS). >gp:gp|D90892|D90892_9 E.co
   li genomic DNA, Kohara clone #446(60.5-60.9 min.). NID: g180
   0074. >gp:gp|AE000353|AE000353_11 Escherichia coli K-12 MG16
   55 section 243 of 400 of the complete genome. NID: g1789037.

   atgaatagaactaaaaatattttagaaattggtacagccattggttatagttcaatgcaa
   ttcgctaatatttcgaaagatatataaatgtaacaacaattgaacgaaatgaagacatgatt
   catcttgcaaaaaagttcataaaaaagtatcgataccagaatcaaatccgtttaattgaa
   tacgatgctttgaatgcatttgaacaagtcaatgacaaacaatatgatatgatatttatc
   gacgcagctaaagcacaatcaatgaaattttttccaactatatacaccgttattaaaaaaa
   ggtggaattgtggttactgataatgtttttatatcatggatttgtttcaaatatagacgtt
   gttcgttcgagaaatgtgaaacaaatggttaagaaagtgcaacagtacaatgaatggttg
   atggagcaatctcaatttacaacaaactttataaatatggatgatggattagcaatatct
   ataaaaggagaatga

Sequence 3040
MNRTKNILEIGTAIGYSSMQFANISKDINVTTIERNEDMIHLAKKFIKKYRYQNQIRLIE
YDALNAFEQVNDKQYDMIFIDAAKAQSMKFFQLYTPLLKKGGIVVTDNVLYHGFVSNIDV
VRSRNVKQMVKKVQQYNEWLMEQSQFTTNFINMDDGLAISIKGE*

Sequence 3041
Contig_0584_pos_9628_10956
is similar to (with p-value 2.0e-23)
   >sp:sp|P75822|YBJT_ECOLI HYPOTHETICAL 53.7 KD PROTEIN IN AR

TP-POXB INTERGENIC REGION.
atgatgccatcagcaaaattaacgcaggctagctttgaagatatggatgcattgttagca
gataattttgcgcgtgcggcgcaaaagcagggtgttaagcatattgtctacatgagtggt
ttaataccagaaatgatgagctatctgcacatttaagaagtcgacttgaatgtgaaaaa
attcttggtgattacggtataccagttagcacattacgtgcaggtttaattattggtgca
aaaggaagttcttatccaattcttaaacgactagttaagagattgccagcaatggttttg
cctagttgggcttacaataaaattgcacctgttgccattgacgatgtgatagatggttta
gcagcattggtgaatcgaacacccaaagataacgaagcaattgatatcacaggtcctgaa
gtgatgaattataaaacgctgatacagcgcacagctaacgtacttgataagcgactgcct
atgcttgatttacctattatacccattatcataagtcggtattgggtacaactgatttca
aatgtaccgaaagaaatggtatatccattaatgaatagtttaactcacgatatggtacca
catcgaaaacgcgttgtgtctaacttgtccgtaggaaatatcacctttgaagatagtgtg
aaaagagcactaagagaagaacaaaagacttctaagaaaaagtcggattcgaaaaattct
caatcatttgggcgtatgcatcaagaaattaaagatgtacgagccattacacggtttaaa
attccggaaggttattcgattaaagatgtgactaaagaatatgcaaaattcatcaataat
atcacactacatctcgttaaaggtacgataaatgaacgagaatttaatatgaatttgccc
ttcattaataaatttattttaaaaatggaacgtgatgaagctgactctacagaagatatg
gtggtatataatattgtgggtggcgatttagcacattcaaatgatggtggaaatgcacgc
tttgaattccgaagaataagaaacaccaatgagggtattattgctttacaagaatatgaa
cctacattaccttgggtagtatataaactaactcaagctaaagcacacaagactgttatg
aatatttttaaaaataaaatggcacgtttatcgcaacaaaaaaatgtgaaagatgaaaca
tatatgtctaatcgtgtaactattggagtaacagtagcatcagcgttcgttattggaagc
gcagtaggggttccaacttttttaaaaagcatcaaatcaaaaagaacacaatgtcgaatgca
gaattataa

Sequence 3042
MMPSAKLTQASFEDMDALLADNFARAAQKQGVKHIVYMSGLIPENDELSAHLRSRLECEK
ILGDYGIPVSTLRAGLIIGAKGSSYPILKRLVKRLPAMVLPSWAYNKIAPVAIDDVIDGL
AALVNRTPKDNEAIDITGPEVMNYKTLIQRTANVLDKRLPMLDLPIIPIIISRYWVQLIS
NVPKEMVYPLMNSLTHDMVPHRKRVVSNLSVGNITFEDSVKRALREEQKTSKKKSDSKNS
QSFGRMHQEIKDVRAITRFKIPEGYSIKDVTKEYAKFINNITLHLVKGTINEREFNMNLP
FINKFILKMERDEADSTEDMVVYNIVGGDLAHSNDGGNARFEFRRIRNTNEGIIALQEYE
PTLPWVVYKLTQAKAHKTVMNIFKNKMARLSQQKNVKDETYMSNRVTIGVTVASAFVIGS
AVGFQLFKKHQIKKNTMSNAEL*

Sequence 3043
Contig_0585_pos_4328_4747
is similar to (with p-value 7.0e-17)
  >sp:sp|P77279|YBBL_ECOLI HYPOTHETICAL ABC TRANSPORTER ATP-B
INDING PROTEIN IN USHA-TESA INTERGENIC REGION. >gp:gp|U82664
|ECU82664_88 Escherichia coli minutes 9 to 11 genomic sequen
ce. NID: g1773084. >gp:gp|AE000155|AE000155_6 Escherichia co
li K-12 MG1655 section 45 of 400 of the complete genome. NID
: g1786692.
atgcaacaaagtgagttaatcggttatacaattgaagataatatgaaatttcctgctgag
gctagaagtgaagctttgaccgtgataaagcgaaacaactcatctctcaagtaggatta
ggtaattatcagttagatgctcaaattgagcacatgtctgggggagagcaacaacgtatt
accatcgctagacaactcatgtatgaacctgaagttttattattggacgaagctactagc
gctttagatacacataataaaaagaaaattgaagaaattatatttaaactagcagataaa
gggattgccattttgtggattacgcatagtgatgaccaaagtatgcgtcattttaagcgt
agaatcacaattactgacggtaagatatcgagtgatgaggagttgaatggtaatgagtaa

Sequence 3044
MQQSELIGYTIEDNMKFPAEARSEAFDRDKAKQLISQVGLGNYQLDAQIEHMSGGEQQRI
TIARQLMYEPEVLLLDEATSALDTHNKKKIEEIIFKLADKGIAILWITHSDDQSMRHFKR
RITITDGKISSDEELNGNE*

Sequence 3045
Contig_0585_pos_5112_5516

is similar to (with p-value 2.0e-32)
  >sp:sp|P77307|YBBM_ECOLI HYPOTHETICAL 28.2 KD PROTEIN IN US
HA-TESA INTERGENIC REGION.
atgttagctaataatggtttaatcgcgattaatctcgcttatcagaatttagaaagagca
tttgttcaagatgtttctgatattgaatccaaacttacgttagcagcgacacctaagctc
gcatcaaaatcagctattagagaaagtatacgcttagcaattgttcctacaattgattct
gtaaaaacatatggtctagtttcaattccaggtatgatgacaggattgattatcggaggc
gttgacccacttcaagcaattaaatttcaattgcttgtcgtgtttattcatacaacagcg
acgattatgtctgcactcattgcaacgtatatgagttacggtcagttctttaatgctcgt
catcaactcattgctagaacgcaacgcacaagacaaagtagttaa

Sequence 3046
MLANNGLIAINLAYQNLERAFVQDVSDIESKLTLAATPKLASKSAIRESIRLAIVPTIDS
VKTYGLVSIPGMMTGLIIGGVDPLQAIKFQLLVVFIHTTATIMSALIATYMSYGQFFNAR
HQLIARTQRTRQSS*

Sequence 3047
Contig_0589_pos_999_559
is similar to (with p-value 1.0e-35)
  >gp:gp|Y10549|BSFMS_1 B.stearothermophilus fms gene. NID: g
2266413.
atgataacaatgaaagatattataagagatggtcatccaacacttcgtgaaaaagcgaaa
gaattaagcttcccactttctaacaatgataaagaaacattgcgcgcaatgcgtgaattt
ctaatcaatagtcaggatgaagaaaccgcaaaacgttatggtttacgttctggcgtaggt
ttagctgctccacaaattaatgaaccaaaacgtatgattgctgtctacttacctgatgat
ggaaacggtaaatcgtatgattatatgctcgtaaatcctaaaataatgagttacagtgta
caagaagcttatttaccaactggcgaaggttgtctaagtgttgatgaaaacatcccaggt
ttagtgcatcgtcatcatttcttaatatcactttcggtttaaaaccacaaaatttaact
atcatagagattcttctttaa

Sequence 3048
MITMKDIIRDGHPTLREKAKELSFPLSNNDKETLRAMREFLINSQDEETAKRYGLRSGVG
LAAPQINEPKRMIAVYLPDDGNGKSYDYMLVNPKIMSYSVQEAYLPTGEGCLSVDENIPG
LVHRHHFLISLSVLKPQNLTIIEILL*

Sequence 3049
Contig_0591_pos_1336_2091
is similar to (with p-value 7.0e-70)
  >sp:sp|Q06753|YACO_BACSU HYPOTHETICAL RRNA METHYLASE IN CYS
S 3'REGION. >gp:gp|D26185|BAC180K_157 B. subtilis DNA, 180 k
ilobase region of replication origin. NID: g467326. >gp:gp|Z
99104|BSUB0001_96 Bacillus subtilis complete genome (section
1 of 21): from 1 to 213080. NID: g2632267.
gtgaatgtggaagatatagtgatagtaggtagacacgcagttaaagaagcaattatatca
ggtcacgccataaataagattttgattcaagacggtataaaaaagcaacaaattaacgac
atttttaaaaaatgcaaaatcacaaaaattaattgtacaaacggtaccaaaatctaaatta
gattttttagcaaatgcacctcaccagggtgtggctgctttagtagcccccatatgaatat
gcaaacttcgatgaattttacaaaaacaaaagaaaaaagcccgttattcaactgttatc
attttagatggtttagaagacccgcataatcttggctctatattaagaacagcagatgct
tctggtgttgatgcggttattatacctaaaagacgatcagttgcgctaacacagaccgtt
gcaaaagcttctacaggagcgattcagcatgttccggttataagggttactaatctttcg
aaaactatcgacgaattaaaagacaacggctttttggattgcggggacagaagctaataat
gcaacggattatagagatttacaagcagatatgtcactaggtattgtaataggtagtgag
gggcaaggtatgagtcgtttagtgagtgataagtgtgattttcatattaagattccaatg
gttggacatgtcaatagcttgaacgcgtctgtggctgcaagtttaatgatgtatgaagta
tatcgtaaacgtcatcagttagaggaaaagtcatga

Sequence 3050
VNVEDIVIVGRHAVKEAIISGHAINKILIQDGIKKQQINDILKNAKSQKLIVQTVPKSKL

DFLANAPHQGVAALVAPYEYANFDEFLQKQKKKARYSTVIILDGLEDPHNLGSILRTADA
SGVDAVIIPKRRSVALTQTVAKASTGAIQHVPVIRVTNLSKTIDELKDNGFWIAGTEANN
ATDYRDLQADMSLGIVIGSEGQGMSRLVSDKCDFHIKIPMVGHVNSLNASVAASLMMYEV
YRKRHQLEEKS*

Sequence 3051
Contig_0592_pos_3596_3057
is similar to (with p-value 3.0e-43)
  >sp:sp|Q47155|DINP_ECOLI DNA-DAMAGE-INDUCIBLE PROTEIN P. >g
p:gp|D38582|ECODINJ_11 Escherichia coli genes for 'YafH, Yaf
I, YafJ, YafK, YafQ, DinJ, YafL, YafM, FhiA, MbhA, DinP, Yaf
N, YafO and YafP. NID: g984576. >gp:gp|D83536|ECOTSF_39 Esch
erichia coli genome, 4.0 - 6.0 min region. NID: g1208942. >g
p:gp|U70214|ECU70214_72 Escherichia coli chromosome minutes
4-6. NID: g1552727. >gp:gp|AE000131|AE000131_10 Escherichia
coli K-12 MG1655 section 21 of 400 of the complete genome. N
ID: g1786415.
atggattatttttttgctcaagttgaaatgagagataatcctaaactaaaaggaaaacct
gtcatcgttggcggtaaagcgagtcatcgaggcgtagtttctacggcatcatacgaagca
agagcttatggtgttcactctgctatgcctatgactcaagcacataagctatgccccaat
ggatattatgtaacaagccgttttgatacttatagagaggtatctggtcaaatcatgaaa
atattcagaagttatacagaattagtagaacccatgtctttagatgaagcttatttagat
attacacatttagtgagaccggatttaccagcatcaaccattgcaaattatattcgcaga
gatatatacgaagtaacacgtttaactgcgtcagctggcgtgtcttataataagttttta
gcaaagttagcgagtggtatgaacaagccgaatggtttgacagtaattgattacaataat
gtacatgaaatattaatgcaattagatattggagattttccaggggtaaaaagcatatag

Sequence 3052
MDYFFAQVEMRDNPKLKGKPVIVGGKASHRGVVSTASYEARAYGVHSAMPMTQAHKLCPN
GYYVTSRFDTYREVSGQIMKIFRSYTELVEPMSLDEAYLDITHLVRPDLPASTIANYIRR
DIYEVTRLTASAGVSYNKFLAKLASGMNKPNGLTVIDYNNVHEILMQLDIGDFPGVKSI*

Sequence 3053
Contig_0593_pos_608_1306
is similar to (with p-value 1.0e-57)
  >sp:sp|Q41364|SOT1_SPIOL 2-OXOGLUTARATE/MALATE TRANSLOCATOR
  PRECURSOR. >gp:gp|A47930|A47930_1 Sequence 1 from Patent WO
9534654. NID: g2301793. >gp:gp|U13238|SOU13238_1 Spinacia ol
eracea envelope membrane 2-oxoglutarate/malate translocator
(SODiT1) mRNA, chloroplast mRNA encoding chloroplast protein
, complete cds. NID: g595680.
atggcctttttcatttcaagaggatttgtaaaaacagggctaggtcgacgtattgctctg
caattcgttaaattatttggaaagaaaacgcttggtttggcttattcacttgttggtgtt
gaccttatcttagctcctgctacgccaagtaatacagcacgtgctggtggtattatgttt
ccaatcattaagtccttgtcagagtcatttggttcatcgccgagagatggttctgagaga
aaaatgggtgcgttttaatctttactgagttccaaggtaatttaattacttcagctatg
tttttaacagctatggccggtaaccctatagcgcaaagtttagctgaaaaaacggcacac
gttcaaattacatggatgaattggttgttgctgctattatacccggattgatttctctc
atcgttgtccctttcattatttataaattataccacctactgttaaagaaacgcctaac
gctaaaaaatgggctactgaacaactagaagaaatgggacatatgtctatagccgaaaaa
ttgatggttggtgtctcttatcatagcattggctttgtgggtattaggaagcttcattaat
gttgatgccacgctcactgcatttattgctttagcattgttactattaacaggtgtatta
gcgtggtcagatattttaaatgaaacagattctgcctaa

Sequence 3054
MAFFISRGFVKTGLGRRIALQFVKLFGKKTLGLAYSLVGVDLILAPATPSNTARAGGIMF
PIIKSLSESFGSSPRDGSERKMGAFLIFTEFQGNLITSAMFLTAMAGNPIAQSLAEKTAH
VQITWMNWFVAAIIPGLISLIVVPFIIYKLYPPTVKETPNAKKWATEQLEEMGHMSIAEK

759

LMVGVFIIALALWVLGSFINVDATLTAFIALALLLLTGVLAWSDILNETDSA*

Sequence 3055
Contig_0593_pos_6260_5115
is similar to (with p-value 4.0e-36)
  >sp:sp|P37469|DNAC_BACSU REPLICATIVE DNA HELICASE (EC 3.6.1
.-). >gp:gp|D26185|BAC180K_4 B. subtilis DNA, 180 kilobase r
egion of replication origin. NID: g467326. >gp:gp|Z99124|BSU
B0021_149 Bacillus subtilis complete genome (section 21 of 2
1): from 3999281 to 4214814. NID: g2636442.
atgttctcacatgatggtatgaagtcatttatggaatatgtattcgaagtcggtaaggta
gatcataacgaaatctatttaaaaaccacaaaagataagtcattcctagatatggacacc
atttcaaatttgtataactcaaaatttataggttacggattctttgaaagatatcaacaa
gatttgctcaatctttatcaaatagagcgtacgcaaaacgtattacaagaattcaattct
gatccgaatatacaaaattttgatgaaatgcttaacaaactacaaaaggtcagtttaatt
agtgcaagtgaagaaagtgggactaaaaaaattgtagatcactttgtcgaagaattatat
agcgaagaaccaaaacaaaaaatcaatacaggttataaactggtggattacaaaataggt
ggtttagaacctacacagttgattgtaatcgctgcgagaccgtcagtaggtaaaacgggg
tttgcgcttaatatgatgcttaatatagcgtctcaaggctataaaacttcattcttcagt
ctagagacaactggcgtgtctgtattgaaaaggatgttatcagcagaaactgggatagaa
ctaactcgtatcaaagaaattaaagatttagaaccggatgaattaacacgtttaacaact
gcagcagacagaatactcaaacttgatatagatatacacgataaaagcaatattactaca
catgatgtacgtaaacaagcgatgaagaacaaagatgtgcaacaggttatcttcattgac
tacttacaacttatgcagacagacagtaagttagatcgtcgtaatggtatcgaaaagata
tcgcgagatttgaagattattgcaaatgaaacaggtgcaattattgtgttgctatctcaa
ttgagcagaggtgtagaaacaagaaatgacaaaagacctatgctatctgacatgaaagaa
gcaggtggaattgaagcagatgcaagtttagctatgttgttatatcgagatgattactac
aaccgtgatgatgttgatgactcaggcaagtcaattgttgaatgtaacatcgcaaagaat
aaagacggagaaacaggtgtagttgagtttgagtactacaagaaaacgcagaggttcttc
acatga

Sequence 3056
MFSHDGMKSFMEYVFEVGKVDHNEIYLKTTKDKSFLDMDTISNLYNSKFIGYGFFERYQQ
DLLNLYQIERTQNVLQEFNSDPNIQNFDEMLNKLQKVSLISASEESGTKKIVDHFVEELY
SEEPKQKINTGYKLVDYKIGGLEPTQLIVIAARPSVGKTGFALNMMLNIASQGYKTSFFS
LETTGVSVLKRMLSAETGIELTRIKEIKDLEPDELTRLTTAADRILKLDIDIHDKSNITT
HDVRKQAMKNKDVQQVIFIDYLQLMQTDSKLDRRNGIEKISRDLKIIANETGAIIVLLSQ
LSRGVETRNDKRPMLSDMKEAGGIEADASLAMLLYRDDYYNRDDVDDSGKSIVECNIAKN
KDGETGVVEFEYYKKTQRFFT*

Sequence 3057
Contig_0594_pos_2706_2050
is similar to (with p-value 2.0e-76)
  >sp:sp|Q59935|MANA_STRMU MANNOSE-6-PHOSPHATE ISOMERASE (EC
5.3.1.8) (PHOSPHOMANNOSE ISOMERASE) (PMI) (PHOSPHOHEXOMUTASE
). >gp:gp|D16594|STRPMI_2 S.mutans pmi gene for mannosephosp
hate isomerase (complete cds) and scrK gene for fructokinase
(partial cds). NID: g451214.
atgagtattttagttattggagcaaatggaggcgtaggttctaaactagtaagtcaatta
aatgaagaacacgttgattttacagctggtgtacgtaaggaagatcaagttaaagaatta
gaaaataaagggattaaagctatattaatagatgtagaaaaaaatagtattaatgattta
aaaaatatctttacagattatgataaagttatcttttcagttggatctggtggaagcact
ggagcggataaaacaatcattgttgatttagatggtgctgtaaaaacaattaaagctagt
aaagaagcgggtatcaaacattatgttatggtatcaacatacgattctagacgtgaagca
ttcgatgcgagtggagatttaaaaccgtatacaattgcaaagcattatgctgatgattac
ttaagaacatcagatcttaattatacaattgtacatccaggttcacttacagatgatgct
ggaactggaaaaatagaagctgatttatatttcgacaaagcaggatcaattccacgtgaa
gatgttgctacagtttttaaaagaagtagtaacttctgatggttttaataatcaagaattc
caaattttaagtggcaatcatggtgttaaagatgcattgaaaaactatgaatcttaa

Sequence 3058
MSILVIGANGGVGSKLVSQLNEEHVDFTAGVRKEDQVKELENKGIKAILIDVEKNSINDL
KNIFTDYDKVIFSVGSGGSTGADKTIIVDLDGAVKTIKASKEAGIKHYVMVSTYDSRREA
5   FDASGDLKPYTIAKHYADDYLRTSDLNYTIVHPGSLTDDAGTGKIEADLYFDKAGSIPRE
DVATVLKEVVTSDGFNNQEFQILSGNHGVKDALKNYES*

Sequence 3059
Contig_0594_pos_1408_470
10   is similar to (with p-value 3.0e-28)
     >sp:sp|Q04304|YMYO_YEAST HYPOTHETICAL 24.9 KD PROTEIN IN RC
A1-NPL6 INTERGENIC REGION. >pir:pir|S54466|S54466 hypothetic
al protein YM9582.15 - yeast (Saccharomyces cerevisiae) >gp:
gp|Z49259|SC9582X_15 S.cerevisiae chromosome XIII cosmid 958
15   2. NID: g807956.
     atgccgttatttttaaaacccattttctggataaagtatggggcagtgataatcttcgt
     caatttgggtatcaactacctaataatcacataggtgaatgttggggaatttcagcacat
     ccacacggaaaaagtgtgattgaaaatggtatatttgctggtcaaacattggatcaagta
     tggaacaatcatagagaaatatttggagattttccaagtaaagattttccattaatggct
20   aaaattgtagatgctgctgcgccattgtctattcatgtacatcccgacgattcttatgca
     tatgaacacgaagaaggtcaatatggaaaatctgaatgttggtacatcattgaagctgat
     gaaggtgcaaagattactataggtacgtatgcgaaatctcgtgatgaatttgaagagcaa
     ttggagcaaggtacatttgaaaattatttgagaacaatacaagtgcaaccaggtgatttt
     tactttataccagctggaacgatacattccataggtgcaggcattatggcgtatgaagtc
25   atgcaatcatcagatatttcatatagaatttatgattatcatagaaaaactgataatagt
     gaggaacgtgaattaaatatagataaggcattagatgttattaattattcaaatgaacta
     cctaatatcactcctcaaaatgaagtgatagaaaatcacaattgtacacatattgtatct
     agtgatttttttactatggttaagtgggatatttctggtactctaaattatatgaagcct
     agagaattttgtcttgtttctgttttagatggacaaggtaaacttattgtagatggtgat
30   atatatgatatatctaaaggttcaaactttgtgttaacttctgaagatttagatagtgtt
     ttcgaaggagattttaaactaatcattagttacatttaa

Sequence 3060
MPLFLKPIFLDKVWGSDNLRQFGYQLPNNHIGECWGISAHPHGKSVIENGIFAGQTLDQV
35   WNNHREIFGDFPSKDFPLMAKIVDAAAPLSIHVHPDDSYAYEHEEGQYGKSECWYIIEAD
EGAKITIGTYAKSRDEFEEQLEQGTFENYLRTIQVQPGDFYFIPAGTIHSIGAGIMAYEV
MQSSDISYRIYDYHRKTDNSEERELNIDKALDVINYSNELPNITPQNEVIENHNCTHIVS
SDFFTMVKWDISGTLNYMKPREFCLVSVLDGQGKLIVDGDIYDISKGSNFVLTSEDLDSV
FEGDFKLIISYI*
40

Sequence 3061
Contig_0595_pos_2494_3402
is similar to (with p-value 2.0e-60)
     >sp:sp|P44331|RBSK_HAEIN RIBOKINASE (EC 2.7.1.15). >pir:pir
45 |B64073|B64073 ribokinase (rbsK) homolog - Haemophilus influ
enzae (strain Rd KW20) >gp:gp|U32732|U32732_6 Haemophilus in
fluenzae Rd section 47 of 163 of the complete genome. NID: g
1573480.
     gtgattgtaattggatcaacaaatgtagataaatttcttaatgttaaaaggtttccaaaa
50   cccggtgagacattacatattaatcaagctcaaaaggagtttggtgggggcaagggagcc
     aatcaagccatagcagctagtagattagcagcagatacaacatttatcagtaaagttggt
     aaagatggcaatgccaactttatattggaagatttcaaaaaagcaggtattcatacacaa
     tatattttaacttcagaaagtgaagaaactgggcaagcatttatcactgttgatgaagca
     ggacaaaatacgattcttgtttacggtggtgcgaatatgacattaagtgcaactgatgtt
55   gagatgagtgcggatgcgtttattggtgcagactttgttgtagcgcagcttgaagttcca
     tttgaggcgatagaacaagcatttaaaattgctcgtaaacaaaatatcactactgtatta
     aatcctgcaccggcaattgaattgcctaagtcactttttagagttaactgatataattatt
     ccaaacgaaacggaagcagaattattaacaggtatttcaatcaataatgaaagtgatatg
     aaagaaacagcaacatattttctcgatttaggtatatctgcagtattaattactttaggg

gagcaaggcacgtattgtgcatatcaagaacaatacaaaatgattcctgcgtgtaatgta
aaagcaatagatacgacagcagcaggagatacatttataggtgcttttttaagtgagtta
aataaagatttgagtaatatagaatcggctattcgacttgcaaatcaagcgtcgtctcta
acggtacaacgaaaaggagcacaagcttctataccaacacgtaaagaagtagaggcagaa

5   tataattaa

Sequence 3062
VIVIGSTNVDKFLNVKRFPKPGETLHINQAQKEFGGGKGANQAIAASRLAADTTFISKVG
KDGNANFILEDFKKAGIHTQYILTSESEETGQAFITVDEAGQNTILVYGGANMTLSATDV

10   EMSADAFIGADFVVAQLEVPFEAIEQAFKIARKQNITTVLNPAPAIELPKSLLELTDIII
PNETEAELLTGISINNESDMKETATYFLDLGISAVLITLGEQGTYCAYQEQYKMIPACNV
KAIDTTAAGDTFIGAFLSELNKDLSNIESAIRLANQASSLTVQRKGAQASIPTRKEVEAE
YN*

15   Sequence 3063
Contig_0596_pos_4550_3993
is similar to (with p-value 3.0e-45)
  >sp:sp|P20368|ADH1_ZYMMO ALCOHOL DEHYDROGENASE I (EC 1.1.1.
  1) (ADH I). >pir:pir|A35260|A35260 alcohol dehydrogenase (EC

20   1.1.1.1) I - Zymomonas mobilis >gp:gp|M32100|ZMOADHA_1 Z.mo
bilis alcohol dehydrogenase I (adhA) gene, complete cds. NID
: g155570.
gtgagtggtatcgaaccaggacaatggttaggcgtatttggtgtgggaggattaggtaat
ttagcattgcaatacgccaaaaacgtaatgggcgcgaaagtcgttgcatttgacattaat

25   gatgataaattaaattttgctaaagagcttggtgctgatgcaatcataaattcaactaat
gttgatcctattgaggaagttaatcgtctaacgaataataaaggattagatgcaacggtg
attactgctgtagctaaaacaccttttaatcaagcagttgatgttgttaaggcgggtgca
cgtgtggtagcagtaggacttccagtagataaaatggatttagatattccacgtttagtg
cttgatggaattgaagtcgttggttcattagttggtaccagacaagatttaagagaagca

30   tttcaatttgctgccgaaaataaagttattcctaaaatccaattaagacaattatctgaa
attaacgatattttttgatgaaatggaaaaaggaacaattacgggtcgaatggtaattgat
atgaaaagcacgcactga

Sequence 3064
35   VSGIEPGQWLGVFGVGGLGNLALQYAKNVMGAKVVAFDINDDKLNFAKELGADAIINSTN
VDPIEEVNRLTNNKGLDATVITAVAKTPFNQAVDVVKAGARVVAVGLPVDKMDLDIPRLV
LDGIEVVGSLVGTRQDLREAFQFAAENKVIPKIQLRQLSEINDIFDEMEKGTITGRMVID
MKSTH*

40   Sequence 3065
Contig_0598_pos_4948_5679
is similar to (with p-value 3.0e-65)
  >sp:sp|P18843|NADE_ECOLI NH(3)-DEPENDENT NAD(+) SYNTHETASE
  (EC 6.3.5.1) (NITROGEN-REGULATORY PROTEIN). >gp:gp|D90817|D9

45   0817_9 E.coli genomic DNA, Kohara clone #326(39.1-39.4 min.)
. NID: g1742837. >gp:gp|D90818|D90818_4 E.coli genomic DNA,
Kohara clone #327(39.2-39.5 min.). NID: g1742849. >gp:gp|AE0
00269|AE000269_3 Escherichia coli K-12 MG1655 section 159 of
  400 of the complete genome. NID: g1788033.

50   atgctaatagataaagcaagatcatttattcagaccatgtatagcgaattaaaatataat
actaatgaaattgaaaatagaatgaaagagattgagcaagaaattaacttgactggtagt
tacacacatacttatgaagaattatcttacggtgcaaaaatggcatggagaaactcaaat
cgttgtattggtagactgttttggaattctttaaatgttaaagatgcccgagatgtatgt
gacgaaaaagaatttataaaatttatacatacacatattaaagaagctactaacggcgga

55   aaaatcaaaccatatattacaatttttagtcctgaagatacacctaaaatttataataat
cagttgattcgttatgctggttatgaaaatgttggcgatccatctgaaaaaaaggttact
cgtttagctgaacatctaggttggaaaggtaaaggttcaaattttgatattttacctctg
atttatcaattgcctaacgacactataaaaatacacgaacttccaaatgatattgtaaaa
gaagtttctatacatcatgaacactatcccaagctttcaaaattaggtttaaaatggtat

gcggtacctattatttcaaatatggatttaaaaatcggtggtattacttaccctacagca
ccttttaatggatggtatatggtaaccgaaattgctgtacgtaatttcacagacacctat
cgtcttatgtaa

5   Sequence 3066
MLIDKARSFIQTMYSELKYNTNEIENRMKEIEQEINLTGSYTHTYEELSYGAKMAWRNSN
RCIGRLFWNSLNVKDARDVCDEKEFIKFIHTHIKEATNGGKIKPYITIFSPEDTPKIYNN
QLIRYAGYENVGDPSEKKVTRLAEHLGWKGKGSNFDILPLIYQLPNDTIKIHELPNDIVK
EVSIHHEHYPKLSKLGLKWYAVPIISNMDLKIGGITYPTAPFNGWYMVTEIAVRNFTDTY
10   RLM*

     Sequence 3067
     Contig_0598_pos_4756_3287
        unknown
15   gtgtaccaatataacgacgatagcttaatgttacacaatgatttatatcaaattaatatg
gctgaaagctactggaatgatggtatccatgaaagaatagcagtgtttgatttgtatttt
cgaaaaatgccatttaatagtggatatgcggtattcaacggattgaaacgcgttgtgaat
ttcatcgaaaactttgggtttacaaatgaagatatcacatatttaaaatcgataggttat
gaagaagattttctaaattacctaaaagatttgaaatttacagggaatattaaatctatg
20   caagaaggtgaaatttgttttggtaatgagccattattaagagttgaagcacctttaatc
caagcgcaacttattgaaactattttgttaaatatcattaatttccaaacattaattgca
actaaagccagccgaattcgtcaaatagcaactcatgacactttgatggaatttggtaca
agaagagctcaagagatcgatgctgcactgtggggcgctagagcagcctttattggaggg
tttgattctacaagtaatgttagagcaggaaaacttttaatatacctgtatctggcaca
25   catgcacacgcactagtacaaacatatggtgatgagtatatagcattcaaaaagtatgct
gagcgacataaaaattgtgtgttcttagttgatacttttcatactttaaaatcaggagta
ccaaccgcaattaaggttgcaaaagagttaggagatactattaattttataggtatcaga
ttagattctggtgatattgcgtacctatctaaagaagctcgtagaatgttagatgaggct
ggttttacagaagctaaaattatcgcatcaaatgatttggatgagcagactattacaagt
30   ttaaaagcacaaggcgctaaagttgacggatggggagtaggtacaaaactgattacagga
tatgatcaaccagccttaggtgcagtttataaattggtttctattgaaacagatgatggc
acaatgagtgatcgcattaaattatcaaataatgctgagaaagttactacaccaggcaaa
aaaaatgtttatcgtattattaataataaaacaggcaaggctgagggcgactatattacg
ctagaaggtgaaaatcctaatgacgaatctccattgaaaatgttccatcctgttcacact
35   tacaaaatgaagtttattaaatcatttaaagcggttaatctacatcaatctatatttgaa
aatggcaaacttgtataccatcttccagatgaatatgaagctcaggactatcttaaaaat
aatttaagtattttatgggaagaaaataaacgatatcttaacccgcaagattatccagta
gatttaagcactaaatgttgggaaaataagcataagcgtattttgaagttgctgaacac
gttaaagagatggaggatgaaaatgagtag
40

     Sequence 3068
VYQYNDDSLMLHNDLYQINMAESYWNDGIHERIAVFDLYFRKMPFNSGYAVFNGLKRVVN
FIENFGFTNEDITYLKSIGYEEDFLNYLKDLKFTGNIKSMQEGEICFGNEPLLRVEAPLI
QAQLIETILLNIINFQTLIATKASRIRQIATHDTLMEFGTRRAQEIDAALWGARAAFIGG
45   FDSTSNVRAGKLFNIPVSGTHAHALVQTYGDEYIAFKKYAERHKNCVFLVDTFHTLKSGV
PTAIKVAKELGDTINFIGIRLDSGDIAYLSKEARRMLDEAGFTEAKIIASNDLDEQTITS
LKAQGAKVDGWGVGTKLITGYDQPALGAVYKLVSIETDDGTMSDRIKLSNNAEKVTTPGK
KNVYRIINNKTGKAEGDYITLEGENPNDESPLKMFHPVHTYKMKFIKSFKAVNLHQSIFE
NGKLVYHLPDEYEAQDYLKNNLSILWEENKRYLNPQDYPVDLSTKCWENKHKRIFEVAEH
50   VKEMEDENE*

     Sequence 3069
     Contig_0598_pos_3057_2467
     is similar to (with p-value 5.0e-55)
55   >pir:pir|A47501|A47501 nitric-oxide synthase (EC 1.14.13.39
     ), endothelial - human >gp:gp|M93718|HUMNIOXSYN_1 Human nitr
     ic oxide synthase mRNA, complete cds. NID: g189211. >gp:gp|L
     10709|HUMNITOX17_1 Human constitutive endothelial nitric oxi
     de synthase gene, exons 25 and 26 and complete cds. NID: g34

8235. >gp:gp|M95296|HUMNOS_1 Human nitric oxide synthase mRN
A, complete cds. NID: g189259. >gp:gp|L26914|HUMNOSA_1 Human
  nitric oxide synthase mRNA, complete cds. NID: g434699.
  gtgaaattaccttatggtgtgcaacaagacgctcatgaagtagaagatgcacttgagttt
5 attaatcctgacacaacatatacagttaatattaaaccagcagttgatcagagtgttcaa
  tcacttagtgaagcaggcatcaaacttactgattttcaaaagggtaatgaaaaagcacgt
  gaacgaatgaaagttcaattttcaattgcttctaatactcaaggtatagtttttaggaact
  gatcactctgccgaaaatattacaggattttacactaaatatggagatggtgctgcggac
  attgcgcctatctttgggttaaataaaagacaaggtaaacaattactagcttatctagga
10 gcacctaaacacctttatgaaaaagtgccaacagctgatttagaagatgataaacctcag
  ttaccagacgaggaagcactaggcgtatcttatcatgatattgatgattatttagaaggt
  aaagaaattcctgcaactgctcgtgaaacaatcgaaaaacattatgttagaaatgcacat
  aagcgtgaacttgcttatacacgatattcatggcctaaatataacaaatga

15 Sequence 3070
  VKLPYGVQQDAHEVEDALEFINPDTTYTVNIKPAVDQSVQSLSEAGIKLTDFQKGNEKAR
  ERMKVQFSIASNTQGIVLGTDHSAENITGFYTKYGDGAADIAPIFGLNKRQGKQLLAYLG
  APKHLYEKVPTADLEDDKPQLPDEEALGVSYHDIDDYLEGKEIPATARETIEKHYVRNAH
  KRELAYTRYSWPKYNK*
20

  Sequence 3071
  Contig_0605_pos_6367_0
  is similar to (with p-value 2.0e-46)
   >sp:sp|Q43157|RPE_SPIOL RIBULOSE-PHOSPHATE 3-EPIMERASE PREC
25 URSOR (EC 5.1.3.1) (PENTOSE-5- PHOSPHATE 3-EPIMERASE) (PPE)
  (RPE). >gp:gp|AF070941|AF070941_1 Spinacia oleracea ribulose
  -phosphate 3-epimerase (RPE) mRNA, nuclear gene encoding chl
  oroplast protein, complete cds. NID: g3264787. >gp:gp|L42328
  |SPIR5P3E_1 Spinacia oleracea nuclear-encoded chloroplast ri
30 bulose-5-phosphate 3-epimerase mRNA, complete cds. NID: g116
  2979.
  atggttaaaattttaccatcacttttatctatagatttttaaatttaaaagaagagctt
  caattgttagaaacagcaaaggtagacggattacactttgatgtaatggacggtaaattt
  gtccctaatatttcaatcggtattccgattttggatgctgttagacaacaatctcatttg
35 ccaatagatgttcatttaatgattgagcaacctgaaaattatattaatctttttgccgaa
  catggtgctgatatgatttctgttcatgttgagtcgacaacacatatacatagagcaatt
  gaacaaattaaacaattagggaaaaaagcaggtgtcgtcatcaatcctggaacatctgta
  gaaacaattttacctatattgagtattgttgattatgttctagtaatgactgtaaatcct
  ggttttggtggacaaacattcatagaacaatgcgtgactaagattgagcaattaaatcaa
40 cttaaacatgaaaatcatttaacttttgatattgaggtagatggaggcattaacgatcaa
  acgagtaaacgatgtgtaga

  Sequence 3072
  MVKILPSLLSIDFLNLKEELQLLETAKVDGLHFDVMDGKFVPNISIGIPILDAVRQQSHL
45 PIDVHLMIEQPENYINLFAEHGADMISVHVESTTHIHRAIEQIKQLGKKAGVVINPGTSV
  ETILPILSIVDYVLVMTVNPGFGGQTFIEQCVTKIEQLNQLKHENHLTFDIEVDGGINDQ
  TSKRCVX

  Sequence 3073
50 Contig_0607_pos_1070_726
  is similar to (with p-value 1.0e-60)
   >pir:pir|I67760|I67760 transposase (insertion sequence IS10
) - Escherichia coli >gp:gp|S67119|S67119_2 BST=somatotropin
  ...BST/beta-Gal fusion protein [Escherichia coli, LBB84, pla
55 smid pXT107, IS10L/R-1, PlasmidTransposonInsertionMutant, 3
  genes, 1679 nt]. NID: g455674.
  atgcagattgaagaaaccttccgagacttgaaaagtcctgcctacggactaggcctacgc
  catagccgaacgagcagctcagagcgttttgatatcatgctgctaatcgccctgatgctt
  caactaacatgttggcttgcgggcgttcatgctcagaaacaaggttgggacaagcacttc

caggctaacacagtcagaaatcgaaacgtactctcaacagttcgcttaggcatggaagtt
ttgcggcattctggctacacaataacaagggaagacttactcgtggctgcaaccctacta
gctcaaaatttattcacacatggttacgctttggggaaattatga

5   Sequence 3074
MQIEETFRDLKSPAYGLGLRHSRTSSSERFDIMLLIALMLQLTCWLAGVHAQKQGWDKHF
QANTVRNRNVLSTVRLGMEVLRHSGYTITREDLLVAATLLAQNLFTHGYALGKL*

     Sequence 3075
10   Contig_0608_pos_9823_8729
     is similar to (with p-value 3.0e-52)
      >sp:sp|Q43157|RPE_SPIOL RIBULOSE-PHOSPHATE 3-EPIMERASE PREC
     URSOR (EC 5.1.3.1) (PENTOSE-5- PHOSPHATE 3-EPIMERASE) (PPE)
     (RPE). >gp:gp|AF070941|AF070941_1 Spinacia oleracea ribulose
15   -phosphate 3-epimerase (RPE) mRNA, nuclear gene encoding chl
     oroplast protein, complete cds. NID: g3264787. >gp:gp|L42328
     |SPIR5P3E_1 Spinacia oleracea nuclear-encoded chloroplast ri
     bulose-5-phosphate 3-epimerase mRNA, complete cds. NID: g116
     2979.
20   atgataactgcagaaaaaaagaagaaaaacaaattcttacctaatttcgaaaaacaatcg
     atctactccttaagatatgacgagatgcaacaatggcttattgatcacggacaacaaaaa
     ttcagagcaaaacaaatttttgaatggttataccaaaagcgtgtgaatactattgatgaa
     atgactaacctgtctaaagagttacgtcaaattctcaaagatcattttgcaatgacgaca
     ttgaccactgttgttaaacaagaaagtaaagatggaacaattaagttcttatttgaatta
25   caagatggttatactattgaaactgttttaatgagacatgaatatggaaattctgtctgt
     gtaacaacacaagtaggatgtagaattggttgtacgttttgtgcttccactttgggcgga
     ttaaagcgtaatttagaggccggagagattgtctctcaagtattaactgtacaaaaggca
     ctagacgaaacgaatgaacgtgtatcacaaattgtcattatgggcataggtgaacctttc
     gagaattatgatgaaatgatggatttcttaagaattgttaatgatgataacagtttaaat
30   attggtgcacgtcatattactgtatctacttcaggaattattccaagaatttatgatttt
     gccgaagaagatatacaaataaattttgctgtgagtcttcatggtgctaaagacgaaata
     agatcaagattaatgcctatcaatcgtgcttataacgttgataagttaatggaagctatt
     cgttattatcaagaaaagacaaatcgccgtgttacttttgaatatggattgtttggtggt
     gttaatgaccaacttgaacatgcgagagatttggcacatttaattaagaatctcaattgc
35   cacgttaatttaataccagttaaccatgtcccagaaagaaattatgtaaagacaccaaaa
     gatgatattttttaaattcgagaaggaattaaagagattaggaattaatgctacaattaga
     cgtgagcaagggtcagatattgatgctgcgtgtggacaattaagagcgaaggaacgacaa
     gtagaaacgaggtaa

40   Sequence 3076
MITAEKKKKNKFLPNFEKQSIYSLRYDEMQQWLIDHGQQKFRAKQIFEWLYQKRVNTIDE
MTNLSKELRQILKDHFAMTTLTTVVKQESKDGTIKFLFELQDGYTIETVLMRHEYGNSVC
VTTQVGCRIGCTFCASTLGGLKRNLEAGEIVSQVLTVQKALDETNERVSQIVIMGIGEPF
ENYDEMMDFLRIVNDDNSLNIGARHITVSTSGIIPRIYDFAEEDIQINFAVSLHGAKDEI
45   RSRLMPINRAYNVDKLMEAIRYYQEKTNRRVTFEYGLFGGVNDQLEHARDLAHLIKNLNC
HVNLIPVNHVPERNYVKTPKDDIFKFEKELKRLGINATIRREQGSDIDAACGQLRAKERQ
VETR*

     Sequence 3077
50   Contig_0608_pos_4794_4123
     is similar to (with p-value 7.0e-63)
      >sp:sp|P36979|YFGB_ECOLI HYPOTHETICAL 43.1 KD PROTEIN IN ND
     K-GCPE INTERGENIC REGION. >gp:gp|D90881|D90881_3 E.coli geno
     mic DNA, Kohara clone #428(56.8-57.0 min.). NID: g1799913. >
55   gp:gp|D90882|D90882_2 E.coli genomic DNA, Kohara clone #429(
     56.9-57.2 min.). NID: g1799919. >gp:gp|U02965|ECU02965_1 Esc
     herichia coli K12 ORF384 gene, complete cds, and ORF337 gene
     , partial cds. NID: g493518. >gp:gp|AE000338|AE000338_3 Esch
     erichia coli K-12 MG1655 section 228 of 400 of the complete

genome. NID: g1788862.
atggttaaaattttaccatcactttttatctatagattttttaaatttaaaagaagagctt
caattgttagaaacagcaaaggtagacggattacactttgatgtaatggacggtaaattt
gtccctaatatttcaatcggtattccgattttggatgctgttagacaacaatctcatttg
ccaatagatgttcatttaatgattgagcaacctgaaaattatattaatcttttttgccgaa
catggtgctgatatgatttctgttcatgttgagtcgacaacacatatacatagagcaatt
gaacaaattaaacaattagggaaaaaagcaggtgtcgtcatcaatcctggaacatctgta
gaaacaatttttacctatattgagtattgttgattatgttctagtaatgactgtaaatcct
ggttttggtggacaaacattcatagaacaatgcgtgactaagattgagcaattaaatcaa
cttaaacatgaaaatcatttaacttttgatattgaggtagatggaggcattaacgatcaa
acgagtaaacgatgtgtagaacagggtgctacaatgttagtcactggttcatacttcttt
aaacaagaggattatgcaaaagtaactagaaactttatgtcatattacagtttttatcata
tatcatttatag

Sequence 3078
MVKILPSLLSIDFLNLKEELQLLETAKVDGLHFDVMDGKFVPNISIGIPILDAVRQQSHL
PIDVHLMIEQPENYINLFAEHGADMISVHVESTTHIHRAIEQIKQLGKKAGVVINPGTSV
ETILPILSIVDYVLVMTVNPGFGGQTFIEQCVTKIEQLNQLKHENHLTFDIEVDGGINDQ
TSKRCVEQGATMLVTGSYFFKQEDYAKVTRNFMSYYSFIIYHL*

Sequence 3079
Contig_0612_pos_14901_13966
is similar to (with p-value 1.0e-57)
>gp:gp|AB009866|AB009866_17 Bacteriophage phi PVL proviral
DNA, complete sequence. NID: g3341907.
atgcacttatctaaaatattaaaacaaggtaaaatcaaagctggagaaccaatggctaaa
acaggtaattcaggtcaatggactactggtccacacgtacacttccaagttgaaagaggt
cgccatgatgacatcacaaacagagggacagtaaaccctgctaaatggctcaaaggtcac
ggtggtggaaaagttggtggtagtggttctgtaaacgcacgtagagcaattcaaagagca
caatctattttaggtggacgttataaatcgtcttatattaccgaacaaatgatgagagtt
gccaaacgtgagtctaacttccaatcagatgcggttaataactgggacatcaacgcacaa
aaaggaacgccttctaaaggtatgttccaaatgattgaaccatcttttagagcatatgct
aaaccaggacacggaaacatcttaaatccaactgacgaagctatatctgctatgcgttac
attgtaggtaagtgggttcctattatggggagttggagaagtgcatttaaacgtgctgga
gattatgcttatgctacaggcggggttattaacactgctggattatataatttggcagaa
gatggataccctgagatagtaatccctacagatccaagcagacaatcagatgcgatgaaa
ttgttacatcttgctgcaagtaaaattagtggaaataacagaaataaacgacctaaccaa
ttacgtacacctaatgttactagtaatacagttgataatgcagaattactactacaaatg
atagaaaatcaacagaaacaaataaacgtgttaatggaaatagcacgaagtaataaaact
attgaaaaacaaccgaaaggttttttcagaacgcgatgtaagtcaggcacaaggttcaagg
ttaagactcgctgcttatagccagggaggtttataa

Sequence 3080
MHLSKILKQGKIKAGEPMAKTGNSGQWTTGPHVHFQVERGRHDDITNRGTVNPAKWLKGH
GGGKVGGSGSVNARRAIQRAQSILGGRYKSSYITEQMMRVAKRESNFQSDAVNNWDINAQ
KGTPSKGMFQMIEPSFRAYAKPGHGNILNPTDEAISAMRYIVGKWVPIMGSWRSAFKRAG
DYAYATGGVINTAGLYNLAEDGYPEIVIPTDPSRQSDAMKLLHLAASKISGNNRNKRPNQ
LRTPNVTSNTVDNAELLLQMIENQQKQINVLMEIARSNKTIEKQPKGFSERDVSQAQGSR
LRLAAYSQGGL*

Sequence 3081
Contig_0612_pos_6018_3844
>pir:pir|S01788|S01788 formate C-acetyltransferase (EC 2.3.
1.54) - Escherichia coli >gp:gp|D90728|D90728_3 Escherichia
coli genomic DNA. (20.3 - 20.7 min). NID: g1651424. >gp:gp|X
08035|ECPFL_1 E. coli pfl gene for pyruvate formate-lyase (E
C 2.3.1.54). NID: g42369. >gp:gp|AE000192|AE000192_6 Escheri
chia coli K-12 MG1655 section 82 of 400 of the complete geno
me. NID: g1787125.

gtgagagaattcatacaattgaactattcattatatgaaggtgacgatgaatttttagaa
ggtcctacaaaagcaactgaaactttatgggatcaagtaatgcaattatcaaaagaagaa
cgtgagcgcggtggcatgtgggacatggatactaaagtggcatcgacaatcacttctcat
gacgctggttacttagacaaagatttagaaaaagttgttggtgttcaaactgaaaaacca
5   ttcaaacgttctatgcaaccattcggtggtattcgtatggcaaaagcagcatgtgaagcg
tatggttacgaattagatccagaaacagaaaaaatcttcactgaatatcgtaaaacacac
aaccaaggtgtattcgatgcatattcaagagaaatgttaaactgtcgtaaagctggtatt
attactggtttgccagatgcttacggacgtggacgtattatcggagactatcgtcgtgtt
gctttatacggtgtagatttcttaatggaacaaaaacttaaagactttaacacaatgtct
10 actgaaatgtctgaagatgtaattcgtttacgtgaagaattatcagagcaatatcgttca
cttcaagatttaaaagaattaggtcaaaaatatggatttgatattagccgtcctgctact
aacttcaaagaagctgtgcaatggttatacttagcatatt°tagctgctatcaaagaacaa
aatggtgcagcaatgagtttaggacgtacttcaacattcttagatatttatgctgaacgt
gattacaaaatggtgacatcactgaacaagaagttcaagaaatcattgaccacttcatt
15 atgaaattgcgtatcgttaaattcgcgcgtacgcctgaatataatgaattattctctgga
gatccaacttgggtaactgaatctatcggtggtgtaggtattgacggacgtccaatggta
actaaaaactcattccgtttcttacactcattagataatttaggtccagccagcaccagca
aacttaacagtgttatggtctactcgcttacctgaaaacttcaaaatctattgtgctaaa
atgagtattaaaacgagctcaatccaatatgaaaatgatgacttaatgcgtgaaagctat
20 ggcgatgattatggtatcgcttgctgtgtatctgccatgaagattggtaaacaaatgcaa
ttcttcggtgcacgtgctaacttagctaaagcattactttacgctatcaatggtggtaaa
gatgaaaaatctggtaaacaagttgggccaagttatgaaggtattaaatcagacgtacta
gattatgatgaagtcttcgaaagatatgaaaaaatgatggactggttagctggcgtatat
atcaactcattaaatatcattcactatgcatgataaatatagctatgaacgtcttgaa
25 atggctttacatgatacagaaattattcgcacaatggcaactggtattgccggattgtct
gtagcagctgactcttatcagcgattaaatatgcacaagttaaacctatccgtaacgaa
gaaggtcttgtaactgactttaaaatcgaaggcgacttccctaaatatggtaataatgac
agtcgtgttgatgaaattgcagtagatttagttgaacgtttcatgactaaattacgtagc
cataaaacataccgtaattctgaacacacaatgagtgtattaacaattacttcaaacgtt
30 gtttatggtaagaaaactggtaacacaccagatggacgtaaagctggcgaaccatttgca
cctggcgcaaacccaatgcatggtcgtgaccaaaaaggtgcattatcttcactaagttca
gtagctaaaataccttacgattgctgtaaagatggtatctcaaatacatttagtatcgta
ccgaaatcactaggtaaagaagaagcagatcaaaataaaaacttaactagtatgttagat
ggttatgcaatgcaacatggtcatcacctcaacattaacgtatttaatagagaaacatta
35 attgatgcaatggaacatccagaagagtatccacaattaacgattcgtgtatctggatac
gctgtaaacttcattaaattaacacgtgaacaacaattagatgttatttcacgtacattc
cacgaatctatgtaa

Sequence 3082
40 VREFIQLNYSLYEGDDEFLEGPTKATETLWDQVMQLSKEERERGGMWDMDTKVASTITSH
DAGYLDKDLEKVVGVQTEKPFKRSMQPFGGIRMAKAACEAYGYELDPETEKIFTEYRKTH
NQGVFDAYSREMLNCRKAGIITGLPDAYGRGRIIGDYRRVALYGVDFLMEQKLKDFNTMS
TEMSEDVIRLREELSEQYRSLQDLKELGQKYGFDISRPATNFKEAVQWLYLAYLAAIKEQ
NGAAMSLGRTSTFLDIYAERDLQNGDITEQEVQEIIDHFIMKLRIVKFARTPEYNELFSG
45 DPTWVTESIGGVGIDGRPMVTKNSFRFLHSLDNLGPAPEPNLTVLWSTRLPENFKIYCAK
MSIKTSSIQYENDDLMRESYGDDYGIACCVSAMKIGKQMQFFGARANLAKALLYAINGGK
DEKSGKQVGPSYEGIKSDVLDYDEVFERYEKMMDWLAGVYINSLNIIHYMHDKYSYERLE
MALHDTEIIRTMATGIAGLSVAADSLSAIKYAQVKPIRNEEGLVTDFKIEGDFPKYGNND
SRVDEIAVDLVERFMTKLRSHKTYRNSEHTMSVLTITSNVVYGKKTGNTPDGRKAGEPFA
50 PGANPMHGRDQKGALSSLSSVAKIPYDCCKDGISNTFSIVPKSLGKEEADQNKNLTSMLD
GYAMQHGHHLNINVFNRETLIDAMEHPEEYPQLTIRVSGYAVNFIKLTREQQLDVISRTF
HESM*

Sequence 3083
55 Contig_0613_pos_918_1343
is similar to (with p-value 1.0e-32)
 >gp:gp|AB009866|AB009866_61 Bacteriophage phi PVL proviral
DNA, complete sequence. NID: g3341907.
atgacaaatacattagaaattaaattattatcagaaaacgcgactatgccgaagagagca

aattctacagatagtggattggacttatacgtatcagaaacgattaacattcctgcacac
gcaactaaagtagttaaaacagatatagcgattaatctgccttatgggtatgaggcgcaa
gtaagacctagatctggtaaatcacttaaaactaaattgcgtgtagcactaggaacaata
gaccaaacataccacaaagaaataggtatcatcacagataacataggtaatgaagatatc
5   acagtagaaaaaggagaaagattagcgcaattagttgtagcgccagttgtatatcctaca
cccaaacaggttgattggtttgaaaatgaaagcgacagaggtgcatatggaagcacagga
gaataa

Sequence 3084
10 MTNTLEIKLLSENATMPKRANSTDSGLDLYVSETINIPAHATKVVKTDIAINLPYGYEAQ
VRPRSGKSLKTKLRVALGTIDQTYHKEIGIITDNIGNEDITVEKGERLAQLVVAPVVYPT
PKQVDWFENESDRGAYGSTGE*

Sequence 3085
15 Contig_0613_pos_2864_3259
is similar to (with p-value 2.0e-21)
>gp:gp|AB009866|AB009866_1 Bacteriophage phi PVL proviral D
NA, complete sequence. NID: g3341907.
gtgaaagaaatgatatataactataagtggatgaccaatattataacatcacaaatttac
20 gatgcagacagtacatctattgcacaatatggtatagaatctgttatgcctaaggctaaa
ggacaagcaggcgataaagtgtttttgaaggttatgaatagaaataaagcgtggagacgt
aatttaaaactcatagagaagattgagtttattgacaagtatgaggagtacataactaac
gagatgaactttcatatattgcagatgattaagttaggtacgaaacttaaaactgttatg
gacttaatggaaataaaaagtaaatcaacattctatggttgtttaaatgaaatagtaaac
25 gtatatatggatgcacagaaaggatattacgattag

Sequence 3086
VKEMIYNYKWMTNIITSQIYDADSTSIAQYGIESVMPKAKGQAGDKVFLKVMNRNKAWRR
NLKLIEKIEFIDKYEEYITNEMNFHILQMIKLGTKLKTVMDLMEIKSKSTFYGCLNEIVN
30 VYMDAQKGYYD*

Sequence 3087
Contig_0613_pos_4377_4844
>gp:gp|X97563|BPHA3GP3_5 Bacteriophage A2 gp3 gene and 4 op
35 en reading frames. NID: g1523807.
atgccgccaagaaaattattatctcaacaaaaaggtaatttaacagtcgaacaacaagaa
aataaagaaaatgcagaaaaagcgatggcgcaactcactgagatagatgaaaaacctcct
gaatggcttgataaagatgcaataaaagaatggcatcgcatattacctttgattcaagaa
ctaccaatagcagctttagatatggggttattagccacctattgtcaaacatatagcaat
40 tacaagaacgccacgatccaattagaaaaagagggtatgatcgtcgaaaccgaaagagga
acgaaattatctagttattacacagtacaaagagatagcgtaaatgctatgaactccatt
tgtcctaaattaggattgacagttgagtcacgtttaaaaatattgtcgccagatacaaag
aaagaaaagaaagatgaatttgaggacttaatgaatggcaaagattag

45 Sequence 3088
MPPRKLLSQQKGNLTVEQQENKENAEKAMAQLTEIDEKPPEWLDKDAIKEWHRILPLIQE
LPIAALDMGLLATYCQTYSNYKNATIQLEKEGMIVETERGTKLSSYYTVQRDSVNAMNSI
CPKLGLTVESRLKILSPDTKKEKKDEFEDLMNGKD*

50 Sequence 3089
Contig_0613_pos_4909_6483
is similar to (with p-value 6.0e-23)
>sp:sp|P14597|DUT_ORFN2 DEOXYURIDINE 5'-TRIPHOSPHATE NUCLEO
TIDOHYDROLASE (EC 3.6.1.23) (DUTPASE) (DUTP PYROPHOSPHATASE)
55 . >gp:gp|M30023|ORFPRTPS_1 Orf virus homologue of retroviral
pseudoprotease gene, complete cds. NID: g332561.
gtgaaagcctgtcaacgccatttagatgacttgaacgattcggaactcccttatcatttt
gatgtaaagaaagctaatcacattattaagtttcttgaaatgttgccagatcctaaaact
ggtaaacaattatcgttaggcggttttcaaaaaattcattgctggtagcttaaatggttgg

```
tacgacagacatgggtacaaaagatttacaaaagcctatatatcaatgagcagaaaaaat
ggtaaaacattattgatctctggaatggcattgtacgatttattgatgggtaaagatccg
ttgaatgaacggttgattggtttgagcgccaattcaagagaccaagctggtatagcatac
gatatgacattggcacaactgaaagctattagaagcgtttctcctaaggttaaatcgatg
```
5
```
actaagataacgccaagtgcaaaagaaatattgaatattaatgatcgaagtaaagttaaa
gccgtttcaaatgaagctgcaaatttagaaggtcatcagtttagctacgcaatcatcgat
gaatatcatgaagctaaagataaaaagatttatgaaacgttaagacgtgggcaagtgcta
ctgcacaaccctatattaattattatctcaacagctggaactaatttgaatggtccgatg
tatgaagaatatttatatattgataagatacttgacggcatagcaaaaaatgaaaactac
```
10
```
tttgtttctgtgctgaacaagatgatgagaaagaagtatatgacgttaaaacttggatt
aaatccaatccacttatggagttgccagaaatggcacaattgttaactaagaatattcaa
ccagaagttaaaactgcaattgatagtggttcaggattaaatgggatattaataaagaat
ttcaatatgtggcgtgcagcaagcacagaatcttatttagatttcaatgattggaagaaa
aatgaaatagactttgatataaatggctctaaaacttatatcggtttagacttatcgcgt
```
15
```
gctgacgacttaaccgcagtatcgttgttcatcttgatgaagataatcaagagtattat
gtaactagtcattcgtttgtggctactaaaggtggattagatggcaagattgatagagac
tttattgattacagacaacttgcagaaagtggttattgtacgattaccgatttacaaagt
ggaattatcaatactgaccaagtttttaaattacattgagaattatatcgaccaatataaa
ttagacgtacaagcgttatgttatgatccttactcaatacatggtgttattgcagaaatt
```
20
```
gagcgtagagattggccttatgatttagtagaaatcagacaagggccacaaacactatct
aatccgatactggattttagactgaaagtgattaatggggacatcaagcatcataaaaat
ccgttactagacattgcagtcaaaaatgctgtggcaaaagataccaatgactcattaatg
attgaaaagaagatgaaccgagaaaaaatagatccactcatggctaccatatttgcttat
gttataaatagctga
```
25

Sequence 3090
```
VKACQRHLDDLNDSELPYHFDVKKANHIIKFLEMLPDPKTGKQLSLGGFQKFIAGSLNGW
YDRHGYKRFTKAYISMSRKNGKTLLISGMALYDLLMGKDPLNERLIGLSANSRDQAGIAY
DMTLAQLKAIRSVSPKVKSMTKITPSAKEILNINDRSKVKAVSNEAANLEGHQFSYAIID
```
30
```
EYHEAKDKKIYETLRRGQVLLHNPILIIISTAGTNLNGPMYEEYLYIDKILDGIAKNENY
FVFCAEQDDEKEVYDVKTWIKSNPLMELPEMAQLLTKNIQPEVKTAIDSGSGLNGILIKN
FNMWRAASTESYLDFNDWKKNEIDFDINGSKTYIGLDLSRADDLTAVSFVHLDEDNQEYY
VTSHSFVATKGGLDGKIDRDFIDYRQLAESGYCTITDLQSGIINTDQVLNYIENYIDQYK
LDVQALCYDPYSIHGVIAEIERRDWPYDLVEIRQGPQTLSNPILDFRLKVINGDIKHHKN
```
35
```
PLLDIAVKNAVAKDTNDSLMIEKKMNREKIDPLMATIFAYVINS*
```

Sequence 3091
Contig_0619_pos_4899_6257
>sp:sp|O69282|MQO_CORGL MALATE:QUINONE OXIDOREDUCTASE (EC 1
40 .1.99.16) (MALATE DEHYDROGENASE (ACCEPTOR)) (MQO).
```
gtgagtcaacccggcgaagaaagttcaaatgtttggaataatgcgggaacaggtcattca
gcattgtgtgaattgaactatacgaaagaaggtaaagatggttcagtagatattactaaa
gcaattcatattaacgagcaatttcaaatatctaaacagttttgggcttatttaatacgt
gaaggtcatattgaaagtccagataaatttattcaatcagtgccacatatgagctttgtt
```
45
```
aaaggggaagaaaatgttaaatttttaaaaagtcgagtggcgagtttacagaaaaatgta
ttatttgaaaaaatgaaaaatttctcaagatccagaaaaaattaactcatgggttccttta
atgatggaaggacgccaatcagatgaagcaattgccattacgtatgacgagacaggtaca
gatgttaactttggtgctttgactaaaaagttaatagctaatttacaacaaaaaaatgtt
ggcattaattataaacatgaagtttttagatataaaaaaattaaataatggtaactggcaa
```
50
```
gttgtggttaaagatttaaatacatcaaatgtaatgaattatgaatctaagttcgtcttc
atcggagctggtggtgcaagtttacctttattacaaaaaacaaagattaaggaatctaaa
cacattggtggtttcccagtaagtggattattttttacgatgtaaaaatccagatgtcata
catagacatcatgcaaaagtctacggtaaagccgaggttggtgcacctccaatgtcagtt
ccacatttagatacacgatttgttaatggtgaaaaatcattactatttggacctttttgca
```
55
```
gggttttcgccaaaattcttaaaaaacggttcatatttagatttagttaaatctgtgaaa
cccaataatatgataacaatgttaagtgctggcgtaaaagaatttaatttgacgaaatat
ttagtttctcaattaatgctttcaaatgaagaacggatcaatgatttgcgtgtattctta
ccagaagcgaaagatgaagattgggaagtaattactgcaggtcaacgtgttcaagtaatt
aaagatacagataagtctaaaggtcaattacaatttggtacggaagtaataacatcagaa
```

```
gatggttcacttgctgcattattaggtgcttcacctggtgcttcgactgctgttgatatc
atgtttgatgtcttgcaacgttgttacaaatcagagtttaagtcatgggaaccaaaaatt
aaagaaatggtcccatcatttggtttaaaattgtcagagcatgaagatatgtaccattca
ataaacgaagaagtaaaaaaatacttgaatgtaaagtag
```

Sequence 3092
VSQPGEESSNVWNNAGTGHSALCELNYTKEGKDGSVDITKAIHINEQFQISKQFWAYLIR
EGHIESPDKFIQSVPHMSFVKGEENVKFLKSRVASLQKNVLFEKMKISQDPEKINSWVPL
MMEGRQSDEAIAITYDETGTDVNFGALTKKLIANLQQKNVGINYKHEVLDIKKLNNGNWQ
VVVKDLNTSNVMNYESKFVFIGAGGASLPLLQKTKIKESKHIGGFPVSGLFLRCKNPDVI
HRHHAKVYGKAEVGAPPMSVPHLDTRFVNGEKSLLFGPFAGFSPKFLKNGSYLDLVKSVK
PNNMITMLSAGVKEFNLTKYLVSQLMLSNEERINDLRVFLPEAKDEDWEVITAGQRVQVI
KDTDKSKGQLQFGTEVITSEDGSLAALLGASPGASTAVDIMFDVLQRCYKSEFKSWEPKI
KEMVPSFGLKLSEHEDMYHSINEEVKKYLNVK*

Sequence 3093
Contig_0620_pos_2802_3113
is similar to (with p-value 7.0e-21)
    >pir:pir|A00461|DEECR NADH dehydrogenase (EC 1.6.99.3) - Es
cherichia coli >gp:gp|D90746|D90746_3 Escherichia coli genom
ic DNA.(24.9 - 25.3 min). NID: g1651543. >gp:gp|V00306|ECNDH
X_1 E. coli gene ndh coding for respiratory NADH dehydrogena
se (a component of the electron transport chain). This enzym
e catalyses the transfer of electrons from NADH to the respi
ratory chain and thus links the major catabolic and energy-p
roducing pathways of the cell. NID: g42112. >gp:gp|AE000211|
AE000211_7 Escherichia coli K-12 MG1655 section 101 of 400 o
f the complete genome. NID: g1787345.
    atgttaaaaaataacgacatgtctgatggttatttaaaaattaaagtaaatggtggagga
tgcacaggattaacttatggtatgtcagccgaagcagaacctggtgaaaatgatgaaatt
ctcgaatactatggtttgaaagttctagtagaccgaaatgatgctcctgtattaaatgga
acaacaattgattttaaacagtcacttatgggtggaggatttcaaataaacaatcctaat
gctattgcctcatgtggttgtggaagttcatttaaaacagctaaagtcgctggaaatcca
gagcaatgttaa

Sequence 3094
MLKNNDMSDGYLKIKVNGGGCTGLTYGMSAEAEPGENDEILEYYGLKVLVDRNDAPVLNG
TTIDFKQSLMGGGFQINNPNAIASCGCGSSFKTAKVAGNPEQC*

Sequence 3095
Contig_0620_pos_3383_4507
is similar to (with p-value 9.0e-18)
    >sp:sp|Q44540|YNIU_AZOVI HYPOTHETICAL 11.0 KD PROTEIN IN NI
FU 5'REGION (ORF6). >gp:gp|M20568|AVINIFC_16 A.vinelandii ma
jor nif gene cluster encoding nitrogen fixation complex, com
plete cds. NID: g758356. .
    atggctcaagatcgtaagaaagtgctcgtattaggtgcaggttatgctggtttacaaact
gtaactaaattacaaaaagaactttctgctgatgcagcggaaattactttaattaataag
aatgaatatcattatgaatcaacttggttgcatgaagcttctgccggtacgattaattat
gaagatttattgtatcctgttgagaaaactgtcaacaaaaataaagtgaattttgttgtt
gctgaggtaacaaaaattgatcgtaatgctaaacgtgtagaaactgataagggtgtttat
gactttgatatcttagttgttgcactaggttttgttagcgaaacatttggtattgatggt
atgaaagaacatgctttccaaattgagaacgtttttaacttctcgtaagttgtctcgtcac
attgaagataagttcgcaaattatgctgcttctaaagaaaaagatgataaagatttatct
attttagttggggggagctggatttacaggaattgaatttctaggtgaattaactgataga
attcctgaattatgcagtaaatatggtgttgatcaaagtaaagtgaagttaacatgtgtt
gaagcagcacctaaaatgttaccgatgttctcagacgacttagttagttatgcagtaaaa
tatttagaagaccgtggagtagaattcaaaattgcaacacctattgtcgcttgtaatgaa
aaaggtttcgttgttgaagtcaatggagaaaaacaacaattagaagccggaacttctgta
```

tggactgctggagtgcgtggaagtcatttaatggaagaatcatttgaaggtgttaaacgt
ggacgtattatcaataaacaagatttaacaattgaaggtcataatgacatctttgttata
ggagattgttcagcgtttattccagctggtgaagagcgtccattaccaacaacagctcaa
attgctatgcaacaaggtgagcatactgctagcaacattaaacgtttattaaatggtgaa
5    tcaacacaagatttccaatatgttaacctaacgatcaagatgcagttgaaaggaaagaga
aagaaactgagtttcaaaaacaacaagatgaagaaattgctttaa


Sequence 3096
MAQDRKKVLVLGAGYAGLQTVTKLQKELSADAAEITLINKNEYHYESTWLHEASAGTINY
10   EDLLYPVEKTVNKNKVNFVVAEVTKIDRNAKRVETDKGVYDFDILVVALGFVSETFGIDG
MKEHAFQIENVLTSRKLSRHIEDKFANYAASKEKDDKDLSILVGGAGFTGIEFLGELTDR
IPELCSKYGVDQSKVKLTCVEAAPKMLPMFSDDLVSYAVKYLEDRGVEFKIATPIVACNE
KGFVVEVNGEKQQLEAGTSVWTAGVRGSHLMEESFEGVKRGRIINKQDLTIEGHNDIFVI
GDCSAFIPAGEERPLPTTAQIAMQQGEHTASNIKRLLNGESTQDFQYVNLTIKMQLKGKR
15   KKLSFKNNKMKKLL*


Sequence 3097
Contig_0620_pos_2198_1134
is similar to (with p-value 7.0e-29)
20    >gp:gp|Y09899|CVPME131_3 C.viguieri phS gene, gene encoding
putative NADH dehydrogenase and two genes encoding unknown
proteins. NID: g2765033.
atgaaaaacttagtattactaggcggggggctatggtaatatgcgaattatgtcgcgcatt
ttacctcattcaattcctgagggatatcacttaactttaatcgaccgcatgccattccac
25   ggtttaaaacctgaattttatgcacttgcagcaggaactaaatctgacaaagaggtgcga
atccaatttccagatagcagtcaaattaatacggtttatggggaaatcagtgatatagat
ttggacgaacaaatgataacagttggaaattcaaaaatagattatgacgaacttatcatt
ggtctagggtgtgaggataaatatcataatgtccctggtgctgaagcatatacacatagc
attcaaacattatctaaatcgcgtgaaacataccatagaattagcgagttacctaaaggg
30   gcacgtgtaggtatcgttggggcaggtttaagtggcattgaattagcgagcgagctacgt
gaaagtcgatcagacttggaaatttttgttatatgatagagggcctcgaatttttaaggaat
tttccagagaaactgagtaaatacatatctaattggttttctaaacacaatgttactgta
gtacctaattcagtcatcgacagagtagaacccggaaaaatttataataatggtaaacca
gaaaatattgatttagtcgtttggacagcaggcatacaacctgttgaaattgtgcgtaat
35   cttcctattgatatgagtaccactggacgcgtaattattaatcagtaccatcaagtccca
acctatagaaatgtttatgtcgtaggtgactgtgctaatttaccacatgcacccagtgct
caactagcagaactacaaggtgaacagattgctgaggtgttgaagaagcaatggaataac
gaaccacttccagataaaatgcctgaaattaaagtacaaggcttttttaggctctttaggt
gacaaacaaggttttgcttatatcatggatcgaacagttaccggacgattagcctctatt
40   ctaaaatcaggtgttctgtggcgctataaatatcataatggttaa


Sequence 3098
MKNLVLLGGGYGNMRIMSRILPHSIPEGYHLTLIDRMPFHGLKPEFYALAAGTKSDKEVR
IQFPDSSQINTVYGEISDIDLDEQMITVGNSKIDYDELIIGLGCEDKYHNVPGAEAYTHS
45   IQTLSKSRETYHRISELPKGARVGIVGAGLSGIELASELRESRSDLEILLYDRGPRILRN
FPEKLSKYISNWFSKHNVTVVPNSVIDRVEPGKIYNNGKPENIDLVVWTAGIQPVEIVRN
LPIDMSTTGRVIINQYHQVPTYRNVYVVGDCANLPHAPSAQLAELQGEQIAEVLKKQWNN
EPLPDKMPEIKVQGFLGSLGDKQGFAYIMDRTVTGRLASILKSGVLWRYKYHNG*


50   Sequence 3099
Contig_0622_pos_4948_4412
is similar to (with p-value 1.0e-61)
>gp:gp|M57622|BACRGB_1 B.stearothermophilus ribosomal prote
in L6 gene, complete cds. NID: g143418.
55   atgagtcgtgttggtaagaaaattattgacattcctagtgacgtaacagtaacttttgac
ggaagtcatgtcactgtaaaaggtccaaaaggtgaattagaaagaactttaaatgaaaga
atgacatttaaacaagaagaaaacactgttgaagttgtaagaccatctgattctaaagaa
gacagaacagatcatggtacaactcgtgctttattaaataatatggtactaggtgtttct
caaggttacgaaaaaacacttgagcttgttggtgtaggttaccgtgcacaaatgcaaggt


771

```
aaagatttagtacttaatgttggatactctcacccagttgaaattaaagcagaagaaggc
attactttcgctgttgagaaaaatacaactgttaaagtatctggtgtttctaaagaacaa
gttggtgcgattgcttctaacattcgttctgtaagacctccagaaccttataaaggtaaa
ggtattcgctaccaaggtgaatatgtacgccgtaaagaaggtaaaactggtaaataa
```

Sequence 3100
MSRVGKKIIDIPSDVTVTFDGSHVTVKGPKGELERTLNERMTFKQEENTVEVVRPSDSKE
DRTDHGTTRALLNNMVLGVSQGYEKTLELVGVGYRAQMQGKDLVLNVGYSHPVEIKAEEG
ITFAVEKNTTVKVSGVSKEQVGAIASNIRSVRPPEPYKGKGIRYQGEYVRRKEGKTGK*

Sequence 3101
Contig_0623_pos_10299_9358
is similar to (with p-value 7.0e-34)
 >gp:gp|AF017231|AF017231_1 Trypanosoma brucei brucei inosin
e-adenosine-guanosine-nucleoside hydrolase mRNA, complete cd
s. NID: g2645494.
atgaccaaagtatattttaatcatgatggtggcgttgatgatctagtgtcactattttta
ttattacaaatggaaaatatagaacttgttggcgtaagtacgattggtgcagactgctat
ttagagccttcattaagtgcttcattaaagattataaatcgtttttcagacgttgaaatt
aatgtagcaccatcatatgaaagagggaaaaatcctttccaaaagaatggagaatgcat
gctttctttatggatgccttaccagtgctaaatgagtcttgtatacccaaaagatgcaag
gctagtgaggacgaggcgtatatagatattattcgtaaagtgaagagttgtgatgagaaa
gttacattgttatttactggaccgcttacagatttagctaaagctataaaatatgacaac
tcaattttaaaaaaatatagagaaattagtttggatgggtggaacgtttttagacaaagga
aatgttgaagaaccagaacatgatggtacagctgaatggaatgcattttgggatcctgag
gctgtaaaagttgtattagatagtgatatgaatgtcgatattgttgctttagaaagtaca
aatcaagtccctctaacaatggaagttcgtcaaatgtgggcagataaaagacaatattta
ggcgttgattttctgggcacaagttacgcagcagtaccaccactcacacattttgtgacc
aattcaacatactttttatgggatgtattaactactgcttatgtgggttctccaaattta
gttgaatcaacgaaattgaaaattgatgtagtcagtcaaggacctagtcaagggagaaca
ttccaatctgaatatggacgtgaagttcaagtcattacggatgtaaataaacaagcattt
tttaactacataacggatttagcaaagaaaatcgagtcctaa
```

Sequence 3102
MTKVYFNHDGGVDDLVSLFLLLQMENIELVGVSTIGADCYLEPSLSASLKIINRFSDVEI
NVAPSYERGKNPFPKEWRMHAFFMDALPVLNESCIPKRCKASEDEAYIDIIRKVKSCDEK
VTLLFTGPLTDLAKAIKYDNSILKNIEKLVWMGGTFLDKGNVEEPEHDGTAEWNAFWDPE
AVKVVLDSDMNVDIVALESTNQVPLTMEVRQMWADKRQYLGVDFLGTSYAAVPPLTHFVT
NSTYFLWDVLTTAYVGSPNLVESTKLKIDVVSQGPSQGRTFQSEYGREVQVITDVNKQAF
FNYITDLAKKIES*

Sequence 3103
Contig_0625_pos_635_1573
is similar to (with p-value 4.0e-70)
 >sp:sp|P54524|YQIG_BACSU PROBABLE NADH-DEPENDENT FLAVIN OXI
DOREDUCTASE YQIG (EC 1.-.-.-). >gp:gp|D84432|BACJH642_230 Ba
cillus subtilis DNA, 283 Kb region containing skin element.
NID: g2627063. >gp:gp|Z99116|BSUB0013_132 Bacillus subtilis
complete genome (section 13 of 21): from 2395261 to 2613730.
NID: g2634723.
```
gtggaaaggggggagtcaaaaggtgaaaacgctaatgattaaagcaatggggacggtgata
cgtttatcgattgagcatcaacatccggatacattacttcaagaagctgaaataaaaatt
cgtgcttgggaatcacaatttagtgctaatgatccgaaatcagatttgatgaatgtgaat
cagcatgcaggtatcgcaccagtcaaggttagttctgagatgtttaacatgatacgtttt
ggttacgaaactacattatcttctaattttaagatgaacattttgatagggccactagtc
aaattatggaaaattggttttaaagatgcattgaaacctaaagaagaggatatacaacgt
gctttattgtgtatggatcctgaaatcttgttctaaattcaaaaacacatgaagtattt
cttacacaatcaggaatggagattgatttaggagctatagttaaaggctattttgctgat
caattacagcaatacttttttagctcatggtgtatcttctggcattatcgatttaggtggt
```

772

```
aatgttttaacaattggtagacaacccgaaacattagaaaaatggcatgtaggtgtacgt
aatccatttcataaggatacactaccactcgttacattaagcgtagagcatcaatcagtt
gtcacatcaggtatctacgaacgctacttcatacaggaaaatcaattatttcatcatata
ttggattcaacaacaggttatcctgtagataatgatatcgctagcgtgacaatcatatct
gatcatgggattgatggcgaggtatggagtacaatttgtagttttggtcagtcacaaaaa
aatattgaattattaaatctcattgacggtattgaaggcattattgtgacaagagatgga
agcgttttaatgacttcgaaaatgcaaaagtatttataa
```

Sequence 3104
VERGSQKVKTLMIKAMGTVIRLSIEHQHPDTLLQEAEIKIRAWESQFSANDPKSDLMNVN
QHAGIAPVKVSSEMFNMIRFGYETTLSSNFKMNILIGPLVKLWKIGFKDALKPKEEDIQR
ALLCMDPENLVLNSKTHEVFLTQSGMEIDLGAIVKGYFADQLQQYFLAHGVSSGIIDLGG
NVLTIGRQPETLEKWHVGVRNPFHKDTLPLVTLSVEHQSVVTSGIYERYFIQENQLFHHI
LDSTTGYPVDNDIASVTIISDHGIDGEVWSTICSFGQSQKNIELLNLIDGIEGIIVTRDG
SVLMTSKMQKYL*

Sequence 3105
Contig_0625_pos_1591_4674
is similar to (with p-value 8.0e-25)
  >sp:sp|P33944|YOJL_ECOLI HYPOTHETICAL 38.5 KD LIPOPROTEIN I
N ADA-OMPC INTERGENIC REGION PRECURSOR. >gp:gp|AE000310|AE00
0310_6 Escherichia coli K-12 MG1655 section 200 of 400 of th
e complete genome. NID: g2367131.

```
atgaagagtattttattttttaagtaataatgtgaaaatattcacaaaaaaactaggagga
tttgctatgagtaaggaaatattcgatactttttaaatttaaatgtggtgccgaattaaaa
aatagagtattaatggcacccatgactatccaagctgggtattttgatggaagtgttaca
tcagaaatgattgattattatcaatttagagctggtgatgcttcagcaatcattgttgaa
agttgttttgttgaaaatcacggacgaggatttccgggagctataggtattgataatgat
gacaaaatacctggactcaaacgtttagcagaagcgattcaagctaagggatcaaaagcg
attttgcaactttatcatgccggaagaatggcaaatcctaaatttaatgaaggagagcag
ccgatatctgcgagccccattgcagcattaagacctgatgctgtaccacctagagaaatg
acacatgctcaaatcaatcagatgattgatgactttggagaggctacacgtcgcgctata
gaagcggggtttgatggtgtcgaaattcatggcgccaacacatacttattacaacaattt
ttctctccacattctaatcggagacaagattcatggggaggcagtcgtgaaaaacgtaca
cgatttccaatcgaagttttgacaaaggttcaacacgtcgttgctgaaaaagaggcttct
cattttattataggatatcgattctcacctgaagaaattgaagaaccaggcatacgtttt
gaagataccatgttttttactaaatacattagcagaatatgaacctgattacttccatata
tcagcaaacagttatcaacgtacatctattgtgaatcaagaagatacagaacctttaatt
aataagtacatcaaaatgcaaagtgcacagttggcaaaaattccattaattggtgtaggt
agtattgcccaacgacaagatgcagaacatgcccttgaactaggatatgatcttttaagt
gttgggaaagcctatttagtggaaccacaatggacagataaaatttcacaaaacgaagaa
gtagaacaatttgtcgatatacatgatcagaaagtacttcacataccatcccccttatgg
aaagtaatggactttatgattttagataaagaagaagagcatcgtaaatatgaaaaatta
aaagcacttcaaaataaaaaagtaaattaacaaaggtacgtatcatgtctatgcaaa
ggtcataatggcaacttacctatgaaagtccaattatcagaagataagattgtaagtatc
gaggtagatgatagcggagagtctgaaggcatagcgaacccagtgtttgaacgtttacct
caagatattatcaatgggcaaacactgaatgtagatgtcatttcaggtgcgacagtaaca
agtgaaggcatcgtgcaaggtattgcagatgcaattgaacaagcaggagaagacccagat
attttacgggcgcgtcctaaaccagtcgttcagtggtctgatgaggttgttgaagagacg
actgatgtcgttgtgattggtacaggaggtgccggactcagtgcagctgctacggcatta
gatgaaggaaaagaagtcatcatgcttgagaaatttgcagctataggtggcaatactatc
cgtacaggtggtcaagtcaacgctgctgagcctaaatggcaaaatgcattcccggcactt
gctggtgaaaaagagacactcatacagttattaaatcatgatgaaaatgatatagatgaa
gcttacattgaagatttcaatactttaaaacgtcaaattaaagactatcttgaaaatagc
agtaatgaagatgaatatcttttttgattctgtcgaattacatcgtattcaaacatattta
ggtggtaaacgtaaagatcgtaataatgtcgaaatttcaggtgattatgatttagttaaa
acactcacagataacgttttggaatcagtatattggttgaaagacaaaggtgtacatttt
gatcgttcgttgtagatatgcctgtgggtgctttatggcgtcgtggtcataaaccaatg
aaagcacaaggtttagagtacattgaaaatttaggagactacgttaaacataatcatggt
```

```
cgtatttttacagaaactactgcagaaaagttaattaaagaaggtaatcaagttgttggt
attgaagcacgtaaagcaaatggtgctaaagtgaagattcatacacgtcatggtgtagtg
ttggctactggtggcctttggagcgaatacaaaaatgctacaacaataatacgtattgg
gataatattcctgatgacattaagacaacgaattctcctgcaattacaggtgatggtata
```
5
```
cgtttaggtgtgcaggcaggcgctgacatcgtaggtatgggattctctcaaatgatgccg
atttctgatccgaagacaggtgcattgtttactggattaattgtaacaccttcaaacttt
gtcttcgttaataaggaaggacagcgttttgttaacgaatttgaaagtagagatgtatta
tctaaggcagcattagaacaaaaagacggtatcttctatattattgcagatgcaaatatt
aaagcactagctatgaatacaactgaggataaaattaatcaagaattagaagacggcact
```
10
```
ttagtaaaagcagataccttagaagcattagcccaaaaattaaacattgatacaactact
tttgtgaacacgattgaaagatataatacccttcgtagaacaaggacaagatgaggatttc
aataaaaatgcatttgatttaaaaattgaaaaagcaccattctacgcgacaccacgtaaa
cctgcaatacatcatactatgggtggtttaaaaataaacacgcatgcacaagttatagat
gttgaaggtcatatcattgaaggtttatatgcggctggtgaagttgccggtggtattcat
```
15
```
gctggtaaccgtttaggcggaaatgcactggcagatattttttacttttggtcgcattgcc
ggtcaaagtgctgtaacgaaataa
```

Sequence 3106
```
MKSILFLSNNVKIFTKKLGGFAMSKEIFDTFKFKCGAELKNRVLMAPMTIQAGYFDGSVT
```
20
```
SEMIDYYQFRAGDASAIIVESCFVENHGRGFPGAIGIDNDDKIPGLKRLAEAIQAKGSKA
ILQLYHAGRMANPKFNEGEQPISASPIAALRPDAVPPREMTHAQINQMIDDFGEATRRAI
EAGFDGVEIHGANTYLLQQFFSPHSNRRQDSWGGSREKRTRFPIEVLTKVQHVVAEKEAS
HFIIGYRFSPEEIEEPGIRFEDTMFLLNTLAEYEPDYFHISANSYQRTSIVNQEDTEPLI
NKYIKMQSAQLAKIPLIGVGSIAQRQDAEHALELGYDLLSVGKAYLVEPQWTDKISQNEE
```
25
```
VEQFVDIHDQKVLHIPSPLWKVMDFMILDKEEEHRKYEKLKALQNKKVKFNKGTYHVYAK
GHNGNLPMKVQLSEDKIVSIEVDDSGESEGIANPVFERLPQDIINGQTLNVDVISGATVT
SEGIVQGIADAIEQAGEDPDILRARPKPVVQWSDEVVEETTDVVVIGTGGAGLSAAATAL
DEGKEVIMLEKFAAIGGNTIRTGGQVNAAEPKWQNAFPALAGEKETLIQLLNHDENDIDE
AYIEDFNTLKRQIKDYLENSSNEDEYLFDSVELHRIQTYLGGKRKDRNNVEISGDYDLVK
```
30
```
TLTDNVLESVYWLKDKGVHFDRSFVDMPVGALWRRGHKPMKAQGLEYIENLGDYVKHNHG
RIFTETTAEKLIKEGNQVVGIEARKANGAKVKIHTRHGVVLATGGFGANTKMLQQYNTYW
DNIPDDIKTTNSPAITGDGIRLGVQAGADIVGMGFSQMMPISDPKTGALFTGLIVTPSNF
VFVNKEGQRFVNEFESRDVLSKAALEQKDGIFYIIADANIKALAMNTTEDKINQELEDGT
LVKADTLEALAQKLNIDTTTFVNTIERYNTFVEQGQDEDFNKNAFDLKIEKAPFYATPRK
```
35
```
PAIHHTMGGLKINTHAQVIDVEGHIIEGLYAAGEVAGGIHAGNRLGGNALADIFTFGRIA
GQSAVTK*
```

Sequence 3107
Contig_0625_pos_7554_4822
40
is similar to (with p-value 3.0e-18)
  >gp:gp|AF061185|AF061185_1 Phytophthora infestans cyst germ
ination specific acidic repeat protein precursor (car90) gen
e, complete cds. NID: g3851513.
```
gtggatgaaatcgttcattatggtggcgaagaaatcaagccaggccataaggatgaattt
```
45
```
gatccaaatgcaccgaaaggtagtcaaacaacgcaaccaggtaagccgggggttaaaaat
cctgatacaggcgaagtagttactccacctgtggatgatgtgacaaaatatggtccagtt
gatggagatccgatcacgtcaacggaagaaattccattcgacaagaaacgtgaattcaat
cctgatttaaaaccaggtgaagagcgtgttaaacaaaaaggtgaaccaggaacaaaaaca
attacaacaccaacaactaagaacccattaacaggggaaaaagttggcgaaggtgaacca
```
50
```
acagaaaaaataacaaacaaccagtagatgaaatcacagaatatggtggcgaagaaatc
aagccaggccataaggatgaatttgatccaaatgcaccgaaaggtagccaagaggacgtt
ccaggtaaaccaggagttaaaaaccctgatacaggcgaagtagtcacaccaccagtggat
gatgtgacaaaatatggtccagttgatggagatccgatcacgtcaacggaagaaattcca
ttcgacaagaaacgtgaattcaatcctgatttaaaaccaggtaaagagcgcgttaaacag
```
55
```
aaaggtgaaccaggaacaaaaacaattacaacaccaacaactaagaacccattaacaggg
gaaaaagttggcgaaggtgaaccaacagaaaaagtaacaaaacaaccagtagatgaaatc
acagaatatggtggcgaagaaatcaagccaggccataaggatgaatttgatccaaatgca
ccgaaaggtagccaagaggacgttccaggtaaaccaggagttaaaaatcctgatacaggc
gaagtagttactccaccagtggatgatgtgacaaaatatggtccagttgatggagatccg
```

```
      attacgtcaacggaagaaattccgtttgataaaaaacgcgaatttgatccaaacttagcg
      ccaggtacagagaaagtcgttcaaaaaggtgaaccaggaacaaaaacaattacaacacca
      acaactaagaacccattaacaggggaaaaagttggcgaaggtgaaccaacagaaaaagta
      acaaaacaaccagtggatgaaatcgttcattatggtggcgaagaaatcaagccaggccat
  5   aaggatgaatttgatccaaatgcaccgaaaggtagccaagaggacgttccaggtaaacca
      ggagttaaaaaccctgatacaggcgaagtagtcacaccaccagtggatgatgtgacaaaa
      tatggtccagttgatggagatccgatcacgtcaacggagaaattccgtttgataaaaaa
      cgcgaatttgatccaaacttagcgccaggtacagagaaagtcgttcaaaaaggtgaacca
      ggaacaaaaacaattacaacaccaacaactaagaacccattaacaggggaaaaagttggc
  10  gaaggtgaaccaacagaaaaagtaacaaaacaaccagtagatgaaatcgttcattatggt
      ggcgaagaaatcaagccaggccataaggatgaatttgatccaaatgcaccgaaaggtagc
      caagaggacgttccaggtaaaccaggagttaaaaatcctgatacaggcgaagtagtcaca
      ccaccagtggatgatgtgacaaaatatggtccagttgatggagatccgattacgtcaacg
      gaagaaattccattcgacaagaaacgtgaattcaatcctgatttaaaaccaggtgaagag
  15  cgtgttaaacagaaaggtgaaccaggaacaaaaacaattacaacaccaacaactaagaac
      ccattaacaggggaaaaagttggcgaaggtgaaccaacagaaaaaataacaaaacaacca
      gtagatgaaatcacagaatatggtggcgaagaaatcaagccaggccataaggatgaattt
      gatccgaacgcaccgaaaggtagccaagaggacgttccaggtaaaccaggagttaaaaat
      cctgatacaggcgaagtagtcacaccaccagtggatgatgtgacaaaatatggtccagtt
  20  gatggagatccgattacgtcaacggaagaaattccgtttgataaaaaacgcgaatttgat
      ccaaacttagcgccaggtacagagaaagtcgttcaaaaaggtgaaccaggaacaaaaaca
      attacaacaccaacaactaagaacccattaacaggagaaaaagttggcgaaggtgaacca
      acagaaaaaataacaaaacaaccagtggatgagatcgttcattatggtggcgaagaaatc
      aagacaggccataaggatgaatttgatccgaacgcaccgaaaggtagtcaaacaacgcaa
  25  ccaggtaagccaggagttaaaaatcctgatacaggcgaagtagtcacaccaccagtggat
      gatgtgacaaaatatggtccagttgatggagatccgattacgtcaacggaagaaattccg
      tttgataaaaaacgcgaatttgatccaaacttagcgccaggtacagagaaagtcgttcaa
      aaaggtgaaccaggaacaaaaacaattacaacgccaacaactaagaacccattaacaggg
      gaaaaagttggtgaaggtgaaccaactctaaagacacctgttaagagtgatgtcagactg
  30  accgcaatattaacaatactcatatctacatga
```

Sequence 3108
VDEIVHYGGEEIKPGHKDEFDPNAPKGSQTTQPGKPGVKNPDTGEVVTPPVDDVTKYGPV
DGDPITSTEEIPFDKKREFNPDLKPGEERVKQKGEPGTKTITTPTTKNPLTGEKVGEGEP
TEKITKQPVDEITEYGGEEIKPGHKDEFDPNAPKGSQEDVPGKPGVKNPDTGEVVTPPVD
DVTKYGPVDGDPITSTEEIPFDKKREFNPDLKPGKERVKQKGEPGTKTITTPTTKNPLTG
EKVGEGEPTEKVTKQPVDEITEYGGEEIKPGHKDEFDPNAPKGSQEDVPGKPGVKNPDTG
EVVTPPVDDVTKYGPVDGDPITSTEEIPFDKKREFDPNLAPGTEKVVQKGEPGTKTITTP
TTKNPLTGEKVGEGEPTEKVTKQPVDEIVHYGGEEIKPGHKDEFDPNAPKGSQEDVPGKP
GVKNPDTGEVVTPPVDDVTKYGPVDGDPITSTEEIPFDKKREFDPNLAPGTEKVVQKGEP
GTKTITTPTTKNPLTGEKVGEGEPTEKVTKQPVDEIVHYGGEEIKPGHKDEFDPNAPKGS
QEDVPGKPGVKNPDTGEVVTPPVDDVTKYGPVDGDPITSTEEIPFDKKREFNPDLKPGEE
RVKQKGEPGTKTITTPTTKNPLTGEKVGEGEPTEKITKQPVDEITEYGGEEIKPGHKDEF
DPNAPKGSQEDVPGKPGVKNPDTGEVVTPPVDDVTKYGPVDGDPITSTEEIPFDKKREFD
PNLAPGTEKVVQKGEPGTKTITTPTTKNPLTGEKVGEGEPTEKITKQPVDEIVHYGGEEI
KTGHKDEFDPNAPKGSQTTQPGKPGVKNPDTGEVVTPPVDDVTKYGPVDGDPITSTEEIP
FDKKREFDPNLAPGTEKVVQKGEPGTKTITTPTTKNPLTGEKVGEGEPTLKTPVKSDVRL
TAILTILIST*

Sequence 3109
Contig_0627_pos_5488_4337
is similar to (with p-value 2.0e-86)
  >sp:sp|P77364|YBBZ_ECOLI HYPOTHETICAL 38.7 KD PROTEIN IN GI
P-FDRA INTERGENIC REGION. >gp:gp|U82664|ECU82664_110 Escheri
chia coli minutes 9 to 11 genomic sequence. NID: g1773084. >
gp:gp|U89279|ECU89279_5 Escherichia coli glyoxylate induced
proteins GlxB1, GlxB2, GlxB3, GlxB4, GlxB6, GlxB7 and GlxB8,
 and glycerate kinase GlxB5 genes, complete cds. NID: g27352
35. >gp:gp|AE000157|AE000157_8 Escherichia coli K-12 MG1655

section 47 of 400 of the complete genome. NID: g1786716.
atgtcaccatataaagttatcattgcacctgattcttttaaagaaagtatgtcggcgaag
gaagctgctttagctattaaagatggattccaagaggtgttcgattccagtacaatatat
gacattattcctatggctgatggtggcgagggaacaaccgaagtattgaaagaagcctta
aatgctacctcttattgtgtagaagtaaaagatccacttaatagaaatatcatggctagt
tatgcgagaagcgacgaacaccaaacagctatcattgaaatggcagctgcttcaggacta
gcattattgagtaaagatgaaagagatccatctattacaacttcgtacggtaccggccaa
ctcattaatgatgcacttaatcacgatgttaataaaattattttaggaataggtggaagt
gccacgaatgatggtggtgtaggaatgttaaaggctttaggtgtctcttttaaagataaa
aacaatcaagagattcgcgatggaggtttagccctatctcaaatagaatacattgatatt
actcgtataaacccacgattgaaagatgtgaatattaaagtagcctgtgatgtaactaat
ccattattaggagacaatggagcaacaatagtttatggtccacaaaaaggcgctcagcaa
aagatgataccaaagttggattcagcattacgtcactatcatgataaaattgaaagagaa
ttaaatatgaatgtaaaagatatcccgggcgctggtgctgcaggaggcatgggaactgca
ttaatcgcgtttctaaacgctaaattacgtcctggaattgatgtagttcttgaagagact
caatttaaacaaaggataaaagatgcgaatttagttgttactggcgaaggtaaaatggat
aaacaaacaatctatggcaaaacacccattggcgtagccaaagttgcaaaatcatatgat
atacccgtcattgctatttgtggtagtttaggaaaagattacgaagcaatttatcaccac
ggtatcgatagcgtgtttagtatcatggaacgtccatgccaccttgacgaagctttgaaa
gaaggcgcacttcatgttaaacatacaacaataaatatcgcacgacttttacaagtaaaa
attgaaaaatga

Sequence 3110
MSPYKVIIAPDSFKESMSAKEAALAIKDGFQEVFDSSTIYDIIPMADGGEGTTEVLKEAL
NATSYCVEVKDPLNRNIMASYARSDEHQTAIIEMAAASGLALLSKDERDPSITTSYGTGQ
LINDALNHDVNKIILGIGGSATNDGGVGMLKALGVSFKDKNNQEIRDGGLALSQIEYIDI
TRINPRLKDVNIKVACDVTNPLLGDNGATIVYGPQKGAQQKMIPKLDSALRHYHDKIERE
LNMNVKDIPGAGAAGGMGTALIAFLNAKLRPGIDVVLEETQFKQRIKDANLVVTGEGKMD
KQTIYGKTPIGVAKVAKSYDIPVIAICGSLGKDYEAIYHHGIDSVFSIMERPCHLDEALK
EGALHVKHTTINIARLLQVKIEK*

Sequence 3111
Contig_0630_pos_4111_3176
is similar to (with p-value 3.0e-25)
  >sp:sp|P19452|HUTG_KLEAE FORMIMINOGLUTAMASE (EC 3.5.3.8) (F
ORMIMINOGLUTAMATE HYDROLASE) (HISTIDINE UTILIZATION PROTEIN
G) (FRAGMENT). >gp:gp|M34604|KPNHUTC_1 K.aerogenes histidine
  utilization repressor C (hutC) gene, complete cds. NID: g14
9203.
atgtatcaacttgcacaatctaatctatggacaggtcgtttagatagtgaaactgatcct
acacaatttagacacttccaaactgttaaattcggtgatttaagtcaattagatttttcg
gatgaacacaaaggcgtgggcttattaggatatgcaattgataaaggagtagaattaaac
aaaggacgtgtaggtgcaaaagaaggtcccaatgccattaagcgagcttttgctggattg
ccagatttgaatcaatgtgaagagattatagattatggtaatgtagaacacaatcatgag
ttgctaatagatacacagcgcgaattcgcagatcttgctgctaagtctatcaaacgacat
aaacaaacatttttacttggtggcggtcatgatatagcatatgcacaatatttagctact
cgtaaagtttatcctgagtcgtcaataggtgtgattaatatagatgcgcactttgacaca
cgcgatgagggttattcaacctctggtatctagttttagacagattctagaagaagatgat
aatgcagattatttagtgttaggtatatctcaaggtggtaatacacaagctttatttaat
tatgctaaagaaaaagatattcaatttgtatatgcgatgaattactacatcaggtatct
ccccccattaaagatatgatagaacgtttatccataatcatgatacggttatgttcaca
atttgcatggatgtagtagatagtgcatttgcaccaggagtcagtgcaccagctgtccta
ggaatatatccacatacagttttcgaacttgctaaacgggtcattccaagtgaaaaagta
aaatctataagtatagctgaaatgaatccgacgtatgattcagatcaaagaactgctaaa
ttagttgctaatttagtacatcattgtttaatttaa

Sequence 3112
MYQLAQSNLWTGRLDSETDPTQFRHFQTVKFGDLSQLDFSDEHKGVGLLGYAIDKGVELN
KGRVGAKEGPNAIKRAFAGLPDLNQCEEIIDYGNVEHNHELLIDTQREFADLAAKSIKRH

KQTFLLGGGHDIAYAQYLATRKVYPESSIGVINIDAHFDTRDEGYSTSGTSFRQILEEDD
NADYLVLGISQGGNTQALFNYAKEKDIQFVYADELLHQVSPPIKDMIERFIHNHDTVMFT
ICMDVVDSAFAPGVSAPAVLGIYPHTVFELAKRVIPSEKVKSISIAEMNPTYDSDQRTAK
LVANLVHHCLI*

Sequence 3113
Contig_0634_pos_2695_2021
  >gp:gp|U40385|SEU40385_1 Staphyloccous epidermidis plasmid
pSK818 insertion sequence IS257(818B) putative transposase g
ene, complete cds. NID: g1762099. >gp:gp|U40386|SEU40386_1 S
taphyloccous epidermidis plasmid pSK818 insertion sequence I
S257(818C) putative transposase gene, complete cds. NID: g17
62101.
  atgaactatttcagatatataaacaatttaacaaggatgttatcactgtagcggttggctac
tatctaagatatgcattgagttatcgtgatatatctgaaatattaagggaacgtggtgta
aacgttcatcattcaacggtctaccgttgggttcaagaatatgccccaattttatatcaa
atttggaagaaaaagcataaaaaagcttattacaaatggcgtattgatgagacgtacatc
aaaataaaaggaaaatggagttatttatatcgtgccattgatgcagagggacatacatta
gatatttggttgcgtaagcaacgagataatcattcagcatatgcgtttatcaaacgtctc
attaaacaatttggtaaacctcaaaaggtaattacagatcaggcaccttcaacgaaggta
gcaatggctaaagtaattaaaggttttaaacttaaacctgactgtcattgtacatcgaaa
tatctgaataacctcattgagcaagatcaccgtcatattaaagtaagaaagacaaggtat
caaagtatcaatacagcaaagaatactttaaaaggtattgaatgtatttacgctctatat
aaaaagaaccgcaggtctcttcagatctacggattttcgccatgccacgaaattagcatc
atgctagcaagttaa

Sequence 3114
MNYFRYKQFNKDVITVAVGYYLRYALSYRDISEILRERGVNVHHSTVYRWVQEYAPILYQ
IWKKHKKAYYKWRIDETYIKIKGKWSYLYRAIDAEGHTLDIWLRKQRDNHSAYAFIKRL
IKQFGKPQKVITDQAPSTKVAMAKVIKGFKLKPDCHCTSKYLNNLIEQDHRHIKVRKTRY
QSINTAKNTLKGIECIYALYKKNRRSLQIYGFSPCHEISIMLAS*

Sequence 3115
Contig_0635_pos_3548_4750
is similar to (with p-value 1.0e-24)
  >gp:gp|U33059|APU33059_5 Actinosynnema pretiosum auranticum
  diaminopimelate decarboxylase (lysA), 3-amino-5-hydroxybenz
oic acid synthase, oxidoreductase, phosphatase, and aminodeh
ydroquinate synthase genes, complete cds; transcription acti
vator gene, partial cds; and unknown genes. NID: g3056877.
  gtgtatcagtcagaaaatcaatcattactttttattgttattttaggttcattaacagca
tttggcccattggctattgatatgtttttacctggactacctaatattagtcatgatttt
gatatttctgcatctacaactcagcttactatctcctttttttatgattggattagcgtta
ggaaatttttggctggccccatatctgatattactggtagaaaaaaaccattaattttc
tcactgattattttttactattgcgagtttaggtattatattcgtcacaaatatatggatt
atgattattttacgatttattcaaggattaactggtggtgcaggtgcagtcatctcaaga
gccattgctagtgatatgtactcaggtaatgcgctaactaaattttttatcattattaatg
cttgtcaatggcattgcgccaattatcgcacctgcgcttggcggtatcattttaaattat
gggccatggcgaattgtatttgtaatactaacaatgtttgggattgtcatgttaatagga
actttatttaaagttcctgagtcgcttgaaaagagcctaagggaaagtagtaacataggt
acgatgctaattaatttcaaagaacttttttaaaacaccccgttttgtattacccatgttg
atacaagggggtgtcatttgtattactatttacttatatttctgcatctccttttatagtt
caaacaatttatggttaacgccattaaacttcagtattatgtttgcttttataggcgtt
acactcattatttcaagccaattaaccggaaaacttgttgactatatagatagattactg
ttgctgagaatcatgtctactatacaagttattggtgttataatcgtatcactaacttta
ctcaaccattggacttttttggatactttcttgtggctttgtgattttagtggcaccagtt
acagggattgcaacgttaggattttcgatagcaatggatgagagtaaaggggccaaaggt
agttcttcaagtttgttgggattgtttcaaactttacttggtggcgtcatctctccactt
gttggtattaagggagacagtaatgcgataccttatataatcgttatcgttattacagca

ataattcttatggttttacagttgattaatgtgaagatatttaaaaaagctaaaattcat
tga

Sequence 3116
VYQSENQSLLFIVILGSLTAFGPLAIDMFLPGLPNISHDFDISASTTQLTISFFMIGLAL
GNFLAGPISDITGRKKPLIFSLIIFTIASLGIIFVTNIWIMIILRFIQGLTGGAGAVISR
AIASDMYSGNALTKFLSLLMLVNGIAPIIAPALGGIILNYGPWRIVFVILTMFGIVMLIG
TLFKVPESLEKSLRESSNIGTMLINFKELFKTPRFVLPMLIQGVSFVLLFTYISASPFIV
QTIYGLTPLNFSIMFAFIGVTLIISSQLTGKLVDYIDRLLLLRIMSTIQVIGVIIVSLTL
LNHWTFWILSCGFVILVAPVTGIATLGFSIAMDESKGAKGSSSSLLGLFQTLLGGVISPL
VGIKGDSNAIPYIIVIVITAIILMVLQLINVKIFKKAKIH*

Sequence 3117
Contig_0635_pos_1170_214
is similar to (with p-value 3.0e-40)
  >sp:sp|P28246|BCR_ECOLI BICYCLOMYCIN RESISTANCE PROTEIN (SU
LFONAMIDE RESISTANCE PROTEIN).
atgttacttttttttatgttttctaattgaattattacttattgttttattatatacgaag
caatcgtttactcttaatttatttagtttttatcttatataccatcattggttttgttatg
atgacttatcatatggtaactgtatcaataccatatgatatgtttatcattgtaattgta
gcaatgatactgttattgattaaatatcgttatattttcaagttgcaaacaggacgtttt
tttattttacaacttagtcatcattttatactgtggggctatttgctgtgagttgttta
tatataagtactattcccctaattatcattaatagcttagctttatgggccgctatcatt
gcatttagtacaatttattcatttatcggatacttatcttggtctacagcttttgaaaat
catcaatattataaacacgtaaagttaattatggtgcttggagctggaattttctctgaa
gaagtgactacgcttcttgctgcaagacttgataaagctttatctgtttatcattcacaa
cggactaaacctatcatcattgtaagtggtggccaaggtcctgacgaaccaatttcagaa
gcacttgcgatgaaaagataccttatagctcacaacgttccggaaaaccatatatttatg
gagaatcaatccacgaatacacgaaccaatttcttatactctaaatctatcattcattcg
atgatgcctacttcaagtcagatgttgtgtgtaacaagtcaatttcatgttttaagagcg
cttaaatttgctaaaaaagctcatctttcttcgatggtattggaagtcgtacaccatac
cactttttggcacaatctatgattatagactttttgggtttaatgtatcaatataaaaca
atacttactatttatttcgctatgttgtttggcttgcaatactacaaaccatataa

Sequence 3118
MLLFLCFLIELLLIVLLYTKQSFTLNLFSFILYTIIGFVMMTYHMVTVSIPYDMFIIVIV
AMILLLIKYRYIFKLQTGRFFILQLSHHFYTVGLFAVSCLYISTIPLIIINSLALWAAII
AFSTIYSFIGYLSWSTAFENHQYYKHVKLIMVLGAGIFSEEVTTLLAARLDKALSVYHSQ
RTKPIIIVSGGQGPDEPISEALAMKRYLIAHNVPENHIFMENQSTNTRTNFLYSKSIIHS
MMPTSSQMLCVTSQFHVLRALKFAKKAHLSFDGIGSRTPYHFLAQSMIIDFLGLMYQYKT
ILTIYFAMLFWLAILQTI*

Sequence 3119
Contig_0637_pos_1621_2709
is similar to (with p-value 9.0e-27)
  >gp:gp|D45211|ARGOD_1 Arthrobacter sp. gene for opine dehyd
rogenase, complete cds. NID: g1060847.
gtgaaaagaggtaaaaaaatgaaaatagctattgtaggttcaggtaatggtgcagtaact
gctgcagtggatatggtagataaaggtcatgatgtacgattatattgtcgtaacgaatct
attagtaaatttgatgtcgccctagaaaaaggtggctttgattttaataatgagggagaa
gagaagtttatagagtttactgatattagtgatgatatggagtatgttttagatggtgca
gacattgttcaggtaattattccttcatcattcattgaatattatgctaaagtgatgtca
aaatttgtgacgaacgaccatctcattttctttaacattgctgcttcaatgggttcaata
cgatttatgaatgtattagaagatcgtcatattgatgtccatccacactttgcagaagca
aatacattaacatatggtacacgtgttgactttaacaatgctaaagtagatttatcttta
aatgttcgtcgggtgttcttttcaacatttgatcgtagtgagttaaatgaaagttatgaa
aaggtatctaaaatttacgattatcttgtaaaagaagaaagtttacttaaaactaatctt
gaaaatggtaacccagaagtacatcctggaccaacattattgaacgttggacgtattgat
tattcagaagagtttttctttatataaagaaggcataacaaaacatactgtgagattatta

```
catgctattgagatagaacgtttaaatttagggagaaaattaggttttgaattatcgact
gccaaagaatcacgtattcaaaggggttatttagaacggaaagacgaggatgaaccgtta
aatcgtcttttttaatactagtcctgtgtttttctcaaattccaggaccgaatcacgttgaa
aaccgttatttaactgaagatatcgcatatggattagtattatggtctagtttaggacgt
```
5   ```
gtcattgatgtcccgacacctaatatcgatgctgttattatgatagcttcaactattctt
gaacgcgatttctttgaagagggcctcactatcgaggaattaggcttagataaattagga
ttagagtaa
```

Sequence 3120
10  VKRGKKMKIAIVGSGNGAVTAAVDMVDKGHDVRLYCRNESISKFDVALEKGGFDFNNEGE
EKFIEFTDISDDMEYVLDGADIVQVIIPSSFIEYYAKVMSKFVTNDHLIFFNIAASMGSI
RFMNVLEDRHIDVHPHFAEANTLTYGTRVDFNNAKVDLSLNVRRVFFSTFDRSELNESYE
KVSKIYDYLVKEESLLKTNLENGNPEVHPGPTLLNVGRIDYSEEFSLYKEGITKHTVRLL
HAIEIERLNLGRKLGFELSTAKESRIQRGYLERKDEDEPLNRLFNTSPVFSQIPGPNHVE
15  NRYLTEDIAYGLVLWSSLGRVIDVPTPNIDAVIMIASTILERDFFEEGLTIEELGLDKLG
LE*

Sequence 3121
Contig_0640_pos_7034_6588
20  is similar to (with p-value 9.0e-16)
 >gp:gp|AB012285|AB012285_1 Photobacterium damsela gene for
sialyltransferase 0160, complete cds. NID: g2988378.
```
atgctaattacgagacaacaacctattgccaaagatttacgtatgatgatggcagcactc
aaaatctctactgactttgaacgaatgggggataatgctgctagtatcgctcatatacgt
```
25  ```
ttaagagttaaaataaatgataactatgtgtttacacgtttaaaaaccatgggtaaatta
gcgatgctcatgttagaagatttaaataacgctattagaaataaagatttaccactgata
aaagaagtcattgagagagatgaagatattgatgatttatacgttaacatcgtcaatacc
agttacttaattgataatgacccattcgtagctggtcaagcacacttagcagctagacac
ttagaacgaataggtgatcatataagcaatattgctgaaagtgtttattattatttaaca
```
30 ```
ggccaacattttgaaactttttgattaa
```

Sequence 3122
MLITRQQPIAKDLRMMMAALKISTDFERMGDNAASIAHIRLRVKINDNYVFTRLKTMGKL
AMLMLEDLNNAIRNKDLPLIKEVIERDEDIDDLYVNIVNTSYLIDNDPFVAGQAHLAARH
35  LERIGDHISNIAESVYYYLTGQHFETFD*

Sequence 3123
Contig_0642_pos_5096_4452
is similar to (with p-value 7.0e-22)
40   >pir:pir|S56598|S56598 yjjG protein - Escherichia coli >gp:
gp|U14003|ECOUW93_286 Escherichia coli K-12 chromosomal regi
on from 92.8 to 00.1 minutes. NID: g1263172. >gp:gp|AE000507
|AE000507_13 Escherichia coli K-12 MG1655 section 397 of 400
of the complete genome. NID: g2367380.
45  ```
gtggatttttatgacgctgagaaaaaagcgtttataatttagcgcagaaatacaatcat
cagccaactcaacaggatttcgaacattttaagaaagtgaaccaagcgcattgggaagca
tttcaacaaaataaattgactaaagatgaagtgctgtcacaacgatttattaattatttt
aatgactatcaaattcatgtaaatggaaaagaagctgatgagtgctttagagctgaatta
gcaaaggcaccagttaaattatttgatcatacattagaagttatacaacaattaaaatta
```
50 ```
aatcattctctctatatagtaaccaatggtgtaacagaaacacagctacgacgaattgct
cagacacaatttaatgaaatatttcaagatgtctttatatctgaacaagctggatttcaa
aagtcgatgacagagttcttcgattttgtgtttgaacatatcggagagaataacaggaat
caaactctaattgtgggagattctttaacgtctgacattttaggtggtaaaaatgctaat
atatcaacatgttggtttaatattagacaaaaagaaaaccatacgtctattcaaccggat
```
55  ```
tatatcattaatgatttatcagaaatgattcgcattgttgagtga
```

Sequence 3124
VDFYDAEKKAFYNLAQKYNHQPTQQDFEHFKKVNQAHWEAFQQNKLTKDEVLSQRFINYF
NDYQIHVNGKEADECFRAELAKAPVKLFDHTLEVIQQLKLNHSLYIVTNGVTETQLRRIA

QTQFNEIFQDVFISEQAGFQKSMTEFFDFVFEHIGENNRNQTLIVGDSLTSDILGGKNAN
ISTCWFNIRQKENHTSIQPDYIINDLSEMIRIVE*

Sequence 3125
5   Contig_0645_pos_2339_3055
    is similar to (with p-value 1.0e-45)
     >gp:gp|AB014075|AB014075_8 Clostridium histolyticum genes f
    or hypoxanthine-guanine phosphoribosyl-transferase (HGPRTase
    ), GTPase and 12 ORFs, complete and partial cds. NID: g38688
10 63.
    atgggacgtaaatggaacaacattaaagagaaaaaagcccaaaaagataaaaatactagt
    agaatatatgccaaatttggtaaagaaatatatgtagctgcaaagtctggtgagcctaat
    ccagagtcaaatcaaactttaagattagtattagaacgtgcaaaaacatattcagtacct
    aatcatattatagatagagctattgataaggctaaaggcgctggtgacgaaaactacgat
15 cacttaagatatgaaggttttggtccgaatggttcaatgcttatagttgacgcattaaca
    aacaatgtaaatcgtacagcatcagatgtacgtgctgcgttcggtaagaatggaggaaat
    atgggagtatctggttcagtagcttatatgtttgaccatactgcaacctttggtgtagaa
    ggtaaatctgtagatgaagtcttagaaacactaatggagcaagatattgatgtaagagat
    gtaattgatgacaatggcttgactattgtttacgcagaaccagatcaatttgcacaagtt
20 caggatgcattacgtgaagctggcgttgaggaatttaaagtagcagagtttgaaatgtta
    cctcaaactgatattgagttgtctgaagaggatcaagctattttttgaaaaattaatcgat
    gcacttgaagacttggaagatgttcaaaatgtttttccataatgtagatttaaaataa

Sequence 3126
25 MGRKWNNIKEKKAQKDKNTSRIYAKFGKEIYVAAKSGEPNPESNQTLRLVLERAKTYSVP
    NHIIDRAIDKAKGAGDENYDHLRYEGFGPNGSMLIVDALTNNVNRTASDVRAAFGKNGGN
    MGVSGSVAYMFDHTATFGVEGKSVDEVLETLMEQDIDVRDVIDDNGLTIVYAEPDQFAQV
    QDALREAGVEEFKVAEFEMLPQTDIELSEEDQAIFEKLIDALEDLEDVQNVFHNVDLK*

30 Sequence 3127
    Contig_0645_pos_4722_5774
    is similar to (with p-value 1.0e-57)
     >sp:sp|P39676|FHP_YEAST FLAVOHEMOPROTEIN (HAEMOGLOBIN-LIKE
    PROTEIN) (FLAVOHEMOGLOBIN).
35 atgtttgaggctaacccagaattattaaacatgtttaaccaaacaaaccaaaagaaaggt
    atgcaatctgctgcattagcacaagcagtactagctgcagcaatgaatattaataattta
    ggtgcaattaagccagcaattatgcctgtggcacataagcactgtgctttacaagtttat
    cccgaacattatccaattgtaggtgaaaatttacttgctgcaattcaagatgtgacaggg
    ttagaaagtgacgacccagtaattcaaacatgggcaaaagcgtatggagaaattgcagat
40 gtatttatcaaattagagcaagaaatttacaaccatatgttatggaaaggttttaaacca
    tttaaaatcacaaacattacacaagaaacaagtgacatcaaatctttcacagttgaatct
    gaagaatatgatttaagtcaattcgaaccaggtcaatacattaccgtagatgtttctagt
    gaaaagttaccatatagagctaaacgtcattattcaatcatagatggagatgaaaatcac
    ttagtatttggtgtcaaacgtgatgtgactactgaacatgaaggtgaagtttcaacaatt
45 ttacatgatgaaatatcagaaggtgacatgattaatttatctgctcctgtaggtggcttt
    tcaatagaaaatactgaaagaccgcaattgtttattggttctggcgtaggtatgacacca
    ttagtttcaatgtttaaaaaagctgcatcattaaacgttccaactcaaatgattcaagcg
    gttgtgacagaagatgaacgaccatttgctcaaaaacttgatagcattacagataattat
    gagcaagcacagctacatttacacgtgaaagataaagaaggttatttagaagctaaagaa
50 ttagaacaatatttaagtgaacagcctgaaatttatatttgtggtggtacgaaattctta
    cattcaattataaattctcttaaagaattaaattatgatatgaatcatgttcatttgaa
    acatttattcctcgtttaagtgttcaagtatag

Sequence 3128
55 MFEANPELLNMFNQTNQKKGMQSAALAQAVLAAAMNINNLGAIKPAIMPVAHKHCALQVY
    PEHYPIVGENLLAAIQDVTGLESDDPVIQTWAKAYGEIADVFIKLEQEIYNHMLWKGFKP
    FKITNITQETSDIKSFTVESEEYDLSQFEPGQYITVDVSSEKLPYRAKRHYSIIDGDENH
    LVFGVKRDVTTEHEGEVSTILHDEISEGDMINLSAPVGGFSIENTERPQLFIGSGVGMTP
    LVSMFKKAASLNVPTQMIQAVVTEDERPFAQKLDSITDNYEQAQLHLHVKDKEGYLEAKE

LEQYLSEQPEIYICGGTKFLHSIINSLKELNYDMNHVHFETFIPRLSVQV*

Sequence 3129
Contig_0650_pos_7844_8518
is similar to (with p-value 1.0e-98)
   >gp:gp|U50335|MSU50335_2 Mycobacterium smegmatis phage resi
stance (mpr) gene, complete cds. NID: g1477566.
atgaactatttcagatataaacaatttaacaaggatgttatcactgtagcggttggctac
tatctaagatatgcattgagttatcgtgatatatctgaaatattaagggaacgtggtgta
aacgttcatcattcaacggtctaccgttgggttcaagaatatgccccaattttatatcaa
atttggaagaaaaagcataaaaaagcttattacaaatggcgtattgatgagacgtacatc
aaaataaaaggaaaatggagctatttatatcgtgccattgatgcagagggacatacatta
gatatttggttgcgtaagcaacgagataatcattcagcatatgcgtttattaaacgtctc
attaaacaatttggtaaacctcaaaaggtaattacagatcaggcaccttcaacgaaggta
gcaatggctaaagtaattaaagcttttaaacttaaacctgactgccattgtacatcgaaa
tatctgaataacctcattgagcaagatcaccgtcatattaaagtaagaaagacaaggtat
caaagtatcaatacagcaaagaatactttaaaaggtattgaatgtatttacgctctatat
aaaaagaaccgcaggtctcttcagatctacggattttcgccatgccacgaaattagcatc
atgctagcaagttaa

Sequence 3130
MNYFRYKQFNKDVITVAVGYYLRYALSYRDISEILRERGVNVHHSTVYRWVQEYAPILYQ
IWKKKHKKAYYKWRIDETYIKIKGKWSYLYRAIDAEGHTLDIWLRKQRDNHSAYAFIKRL
IKQFGKPQKVITDQAPSTKVAMAKVIKAFKLKPDCHCTSKYLNNLIEQDHRHIKVRKTRY
QSINTAKNTLKGIECIYALYKKNRRSLQIYGFSPCHEISIMLAS*

Sequence 3131
Contig_0650_pos_6596_5145
   >gp:gp|U40385|SEU40385_1 Staphyloccous epidermidis plasmid
pSK818 insertion sequence IS257(818B) putative transposase g
ene, complete cds. NID: g1762099. >gp:gp|U40386|SEU40386_1 S
taphyloccous epidermidis plasmid pSK818 insertion sequence I
S257(818C) putative transposase gene, complete cds. NID: g17
62101.
atggttgaaaaattaaaacatgaatggtttaaccagccaggtaaaaatatacttgccggt
atcgtggttgccttagctttaatccctgaagctatcgcattttcaattattgctggcgta
gatccaatggttggttttgtatgcttcatttatcatcgctgttgttactgctattgttggt
ggtagacctgcaatgatatcaggtgcaacaggggctgttgccttattagttacaccactt
gtgaaagattatggtgtagaatatctttttagctgccacgatattaatgggagtaattcaa
ttagttttaggccttctcaaagtggggcgtttaatgaaatttatacctcattccgtcatg
ataggttttgtaaatgcattaggtattatgattttcatgtcccaaatagaacatatcttc
ggtatttcaatctcaacttatatatatgtaattgtaacattactcattgtatatataatc
cctaaattttttcaaagcaataacctgcaccattaatagctatcatcgtattgactgctctt
tatatgtatacaggttctgacgtgagaactgttggtgatttaggaaatattaagcaagct
ttaccgcacttttttaattcctaatgttcccttttaatttagaaacacttcaaatcattttt
ccatactcgctatctatggctattgtaggtctagtagaaagtttacttactgctaaaatt
gtagatgatgcaacagacacttatagtagtaaaaacagagaatctcgtggccaaggcatt
gctaatatgattacaggattattcggtggtatgggaggttgtgccatgattggacaatct
gtaatcaatgtcaaatcaggtgcaaacagtagattatctactttttctgctggtgttgtc
ttaatattcatgattattgttcttggaggacttgttgttcaaattccaatgccaattttta
gcaggtattatggttatggtttcgattggtacatttgattggaattcttttaaatatatt
caaaaagcaccaaaaacagatgcagttgttatgatacttacagtgataattgtactgatg
acacataacttagctctcggcgtggtcgtaggtgttattttcagtgctttattcttttgct
actaaaatatcaaaagtagaagtaacatctgagaagtttggtaaaactaaccgtttatct
tttaaaggtcaaatcttttttgtttctattgactctatgatggatcaaattagctttaat
attgaaaatagtattatagaattaaactttaataatgctcatttatgggatgattcagca
gtagatgctattgatacaatggtaaggaagttcgaagaaaaaaataacattgttcatgta
gaaaaactaaattcagatagtcgtaaaatagtctcagaattaagcaaactaaatgaaaat
catttaaactaa

Sequence 3132
MVEKLKHEWFNQPGKNILAGIVVALALIPEAIAFSIIAGVDPMVGLYASFIIAVVTAIVG
GRPAMISGATGAVALLVTPLVKDYGVEYLLAATILMGVIQLVLGLLKVGRLMKFIPHSVM
IGFVNALGIMIFMSQIEHIFGISISTYIYVIVTLLIVYIIPKFFKAIPAPLIAIIVLTAL
YMYTGSDVRTVGDLGNIKQALPHFLIPNVPFNLETLQIIFPYSLSMAIVGLVESLLTAKI
VDDATDTYSSKNRESRGQGIANMITGLFGGMGGCAMIGQSVINVKSGANSRLSTFSAGVV
LIFMIIVLGGLVVQIPMPILAGIMVMVSIGTFDWNSFKYIQKAPKTDAVVMILTVIIVLM
THNLALGVVVGVIFSALFFATKISKVEVTSEKFGKTNRLSFKGQIFFVSIDSMMDQISFN
IENSIIELNFNNAHLWDDSAVDAIDTMVRKFEEKNNIVHVEKLNSDSRKIVSELSKLNEN
HLN*

Sequence 3133
Contig_0651_pos_3288_1672
is similar to (with p-value 5.0e-34)
  >sp:sp|P31679|YAAU_ECOLI HYPOTHETICAL METABOLITE TRANSPORT
  PROTEIN IN CARB-KEFC INTERGENIC REGION (ORF65/66). >gp:gp|AE
  000114|AE000114_12 Escherichia coli K-12 MG1655 section 4 of
  400 of the complete genome. NID: g1786217.
atgaaaataagggatgaaaatatggatttcgttaaatcaaaaactgacttatttagactc
atagacaatgaagcgcaaacatcgacatctaagatggttttattcttaatattaggaact
atattttttagatgcatatgatattactattttaggtacaatgactgatcaactcactcaa
cagtttcacttatcaccatcaacgctatctatagtaatgacctctttacctattggtgca
ttatttggtgcattacttggtggtacattagcacatcagtttggacgcaagcatatttta
tcaattgccttactaacactcactgtaacctctcttggtgcggcactcgcaccaaatgta
attattctaataatatgtcgttgtataatggggtttgctattggaatggatagtccagtt
gctttcacttttattgcggaaataagtaatttaaagcacaaaggaagaaatgttaactat
tggcaagtcgtttggtatgttgcaatagttacttctgctttagtggtcattgcgttcttt
atgctaggggctggtgcacatttgtggagatatgcaattggatttggtgcacttattgct
tttgtcttgtacatcttgcgaattaaatatttacacgaaagtccgacatgggtgattaat
cattattctttagaaaaagcaactgaatttgtaagaaaatattatcataaagacatccac
ctagagggaacgcttgaagatgatttaagttctgatgtgacttcgccacataattcttgg
acagacttatttaaaccgagatatataaaaagaattatcctggcgactgcgatttcaaca
ttacaaggtatgcagtactatggtgtcggtttatacatacctattattgcaacttatctt
attagtaaggataaaataggtgtattattaggtactgctatagtcaatatagcaggtatt
ctaggcgcatatttaggtgctcaattgacttataaattaggtacacgcaagcttacaatg
ataggcttcacacttgtattactttcaatggtatgtgtaggactcttttatcatcatcta
ccaatgcttcttaacactttccttattggattatttttatttggccattcaggaggtcct
ggtactcaaggaaaaacaattggtgccttatcattcccgactcatttacgttcacaagct
actggctttgtagaatctgtaagtcgtactggtagtatcataggtacttttgtctttcca
atcattcttgctgcagtaggtctaacgaatactatgttaatcttgtccattgtcccttta
ctcggaattatcataacagtatctataaaatgggaagctgtcggtaaggaatacattgtt
gaatataacgctactttggcattaaacgatatagaaagatcgataattagaaaagaatta
acattagcttttaaaagaagcgaagtcaaactcagtcgaatggaaagacgtatcattcga
ttactacttaatgattacaagccaaggagattgctatggtttttaaatttggaatccaaa
gttgtttataatgcgattcaacgtagtaaatgtaaacttaaaagaagttttgaataa

Sequence 3134
MKIRDENMDFVKSKTDLFRLIDNEAQTSTSKMVLFLILGTIFLDAYDITILGTMTDQLTQ
QFHLSPSTLSIVMTSLPIGALFGALLGGTLAHQFGRKHILSIALLTLTVTSLGAALAPNV
IILIICRCIMGFAIGMDSPVAFTFIAEISNLKHKGRNVNYWQVVWYVAIVTSALVVIAFF
MLGAGAHLWRYAIGFGALIAFVLYILRIKYLHESPTWVINHYSLEKATEFVRKYYHKDIH
LEGTLEDDLSSDVTSPHNSWTDLFKPRYIKRIILATAISTLQGMQYYGVGLYIPIIATYL
ISKDKIGVLLGTAIVNIAGILGAYLGAQLTYKLGTRKLTMIGFTLVLLSMVCVGLFYHHL
PMLLNTFLIGLFLFGHSGGPGTQGKTIGALSFPTHLRSQATGFVESVSRTGSIIGTFVFP
IILAAVGLTNTMLILSIVPLLGIIITVSIKWEAVGKEYIVEYNATLALNDIERSIIRKEL
TLAFKRSEVKLSRMERRIIRLLLNDYKPKEIAMVLNLESKVVYNAIQRSKCKLKRSFE*

Sequence 3135

Contig_0652_pos_1335_2843
>sp:sp|069282|MQO_CORGL MALATE:QUINONE OXIDOREDUCTASE (EC 1
.1.99.16) (MALATE DEHYDROGENASE (ACCEPTOR)) (MQO).
gtgagcaaaaaaatggctaataaagagtcaaaaaatgttgttattattggcgctggtgtc
ttaagtacgacatttggttctatgattaaagaattagaacctgattggaacatcaaactc
tatgaacgcttagatcgtccaggtattgaaagttctaacgaaagaaacaatgccggtaca
ggacatgcggcgttatgtgaattgaactatacagtacaacaacctgatggttcaattgat
atagaaaaagccaaagaaatcaacgaacaattcgagatttcaaaacaattctggggtcac
ttagtaaaaagtggtaacatcagtaaccctagagatttcattaatccacttcctcacatt
agtttcgtaagaggtaaaaataacgttaaattcttaaaaaaccgttacgaagcaatgcgt
aacttccctatgttcgataacatcgaatatacagaagatatcgaagaaatgagaaaatgg
atgccattaatgatgacaggccgtactggtaacgaaatcatggcggctagtaaaatcgac
gaaggtacagatgttaactacggtgaattaactcgtaaaatggcaaaaagtattgaaaaa
catccaaatgctgatgttcaatacaaccacgaagtaattaatttcaatcgtcgtaaagac
ggtatttgggaagttaaagttaaaaaccgtaattctggagacgttgaaactgttctagct
gattatgtatttatcggtgcaggcggtggcgctattccactattacaaaaaactggtatc
ccagaaagtaaacatcttggtggattccctatcagtggtcagttcttaatttgtacaaac
cctgatgtaattaatgaacatgacgtcaaagtatatggtaaagaaccaccaggcacacct
ccaatgactgtaccacatttagatacacgttatatcgatggtgaaagaacattattattt
ggaccatttgcaaatattggccctaaattcttaagaaacggttctaacttagacttattc
aaatcagttaaaccttataacatcacaacattactagcatctgcagttaaaaacttacct
ttaatcaaatactctatcgaccaagtattaatgactaaagaaggttgtatgaaccatcta
cgcacgttctaccctgaagctcgtgacgaagattggcaattatacactgcaggtaaacgt
gttcaagttatcaaagatactaaagaacacggtaaaggattcattcaattggtacagaa
gttgttaactctaaagaccactctgttatcgcactattgggtgaatcacctggagcatca
acttcagtatcagtagccctagaagtttttagagaaaaactttgctgagtatgaaaaagat
tggactccaaaattacaaaaaatgatcccatcatatggtaaatctcttatcgatgatgtt
aagttaatgagagcaactcgtaaacaaacatctaaagatttagaattaaattattacgaa
tctaaataa

Sequence 3136
VSKKMANKESKNVVIIGAGVLSTTFGSMIKELEPDWNIKLYERLDRPGIESSNERNNAGT
GHAALCELNYTVQQPDGSIDIEKAKEINEQFEISKQFWGHLVKSGNISNPRDFINPLPHI
SFVRGKNNVKFLKNRYEAMRNFPMFDNIEYTEDIEEMRKWMPLMMTGRTGNEIMAASKID
EGTDVNYGELTRKMAKSIEKHPNADVQYNHEVINFNRRKDGIWEVKVKNRNSGDVETVLA
DYVFIGAGGGAIPLLQKTGIPESKHLGGFPISGQFLICTNPDVINEHDVKVYGKEPPGTP
PMTVPHLDTRYIDGERTLLFGPFANIGPKFLRNGSNLDLFKSVKPYNITTLLASAVKNLP
LIKYSIDQVLMTKEGCMNHLRTFYPEARDEDWQLYTAGKRVQVIKDTKEHGKGFIQFGTE
VVNSKDHSVIALLGESPGASTSVSVALEVLEKNFAEYEKDWTPKLQKMIPSYGKSLIDDV
KLMRATRKQTSKDLELNYYESK*

Sequence 3137
Contig_0652_pos_7388_7999
is similar to (with p-value 9.0e-32)
>sp:sp|P15029|FECD_ECOLI IRON(III) DICITRATE TRANSPORT SYST
EM PERMEASE PROTEIN FECD. >pir:pir|S56513|S56513 citrate-dep
endent iron transport protein fecD - Escherichia coli >gp:gp
|U14003|ECOUW93_200 Escherichia coli K-12 chromosomal region
from 92.8 to 00.1 minutes. NID: g1263172. >gp:gp|AE000499|A
E000499_8 Escherichia coli K-12 MG1655 section 389 of 400 of
the complete genome. NID: g1790732.
atgatcataatgatatttccatcagcacctctatttgttcttcctttaggttcatttatc
ggtgctttgacaataagtattattctttcagttcttatttcaaaatttgatgtaaaggga
tcaaaattagcattgataggtttagcgataggtgcaatttgtacggccattgtccaattc
ttgcttatacgtaatcctcttgatgcaaataatgcgttattatggttgactggtagttta
tacggtcataatatagtcaatttttatagtttattaccatggtttattatcactgtacct
atagtattgttattagggtatcaacttgatattttaaatttaggtgatcatgtagccatt
gcactaggagcacgtgtaaaaatcttaaaaatgattttacttgtattagcagtaatgtta
gcaggtgcttccattgcggtagtaggggggtattagttttttaggtcttatagcacctcat

attgcacgtcaacttgtcggccataaaaatatacatgttataatcatgtcaggtttggta
ggagcaatattattaacttttggtgatggtttagcaagaggtatacaacctcctcttgat
attcctgtatga

Sequence 3138
MIIMIFPSAPLFVLPLGSFIGALTISIILSVLISKFDVKGSKLALIGLAIGAICTAIVQF
LLIRNPLDANNALLWLTGSLYGHNIVNFYSLLPWFIITVPIVLLLGYQLDILNLGDHVAI
ALGARVKILKMILLVLAVMLAGASIAVVGGISFLGLIAPHIARQLVGHKNIHVIIMSGLV
GAILLTFGDGLARGIQPPLDIPV*

Sequence 3139
Contig_0653_pos_2824_2252
is similar to (with p-value 6.0e-31)
 >sp:sp|P14776|DHSS_SYNP1 SOLUBLE HYDROGENASE, SMALL SUBUNIT
 (EC 1.12.-.-) (TRITIUM EXCHANGE SUBUNIT). >pir:pir|S06919|H
QYCSS soluble hydrogenase (EC 1.12.-.-) small chain - Synech
ococcus sp. (PCC 6716) >gp:gp|X16658|SYNSOLHY_1 Synechococcu
s DNA for the small subunit of soluble hydrogenase. NID: g48
053.
atgctaccaccaggtctagcatttgttgcttatagcgatagagcaaaaaaacgatttgct
gatgtaaaaacaccgagattctatttagatttaaataaatacataaaatcacaagagcaa
aattcaacgcccttcacccctaatgttggtctatttagaggaataaatgcttatgtagaa
cttgtaaaaaaagaaggattaaatcacgttatttcacgccattttaaaatacgtaatgcc
ttaagagcagcactaaaggcacttgaattagaattattagtaaaagatgatgctcatgcc
tcacctactgtgacctcatttgttccaaaaaatcaagaagaacttaatatcattaaaaat
caacttaaatctcaattcaatataactattgctgggggacaaggacacttaaaaggacaa
attttgagaattggtcacatggggaaaatatctccttttgatattttagcagtcgtgtct
gcattggaaattattttaacttctaatagaaatgtcaattatattggaacagggataact
caatttatggaggttattagacatgagtcataa

Sequence 3140
MLPPGLAFVAYSDRAKKRFADVKTPRFYLDLNKYIKSQEQNSTPFTPNVGLFRGINAYVE
LVKKEGLNHVISRHFKIRNALRAALKALELELLVKDDAHASPTVTSFVPKNQEELNIIKN
QLKSQFNITIAGGQGHLKGQILRIGHMGKISPFDILAVVSALEIILTSNRNVNYIGTGIT
QFMEVIRHES*

Sequence 3141
Contig_0657_pos_4668_5093
is similar to (with p-value 2.0e-19)
 >gp:gp|AF001974|AF001974_1 Thermoanaerobacter ethanolicus p
utative TrkG gene, partial cds, and putative TrkA, xylose is
omerase (xylA) and xylulose kinase (xylB) genes, complete cd
s. NID: g2581794.
gtgagaccgactgtaccgaatgctgatacgacttcgaagagtattttgattaaaggtata
ttggaattgattatggtaagtataaacgtaaccataccgataaatgcaatagaaatgaga
atagttacgaaagacaactgtatgtatctttctgatatttctctattaaatatagagttg
tttttttctttgcgtatcgtattaaaaatagcgattgttgcaataacaaaagttgtaacc
tttataccacctgcagcactcaatggtgcacctccaataaacatgagagccataagtaat
aaagctgtcggtgtttttaatgtttccaacgtcaattgtgttaaatcctgcagtccttgtt
gtcactgattggaaaaatgcatttcctatttttttcaattaatcccatgtgtaacatagag
ttttga

Sequence 3142
VRPTVPNADTTSKSILIKGILELIMVSINVTIPINAIEMRIVTKDNCMYLSDISLLNIEL
FFSLRIVLKIAIVAITKVVTFIPPAALNGAPPINMRAISNKAVGVLMFPTSIVLNPAVLV
VTDWKNAFPIFSINPMCNIEF*

Sequence 3143
Contig_0660_pos_3013_3495

is similar to (with p-value 3.0e-18)
  >gp:gp|D89592|D89592_3 Vibrio alginolyticus rhlE, KtrA and
  KtrB genes, complete cds. NID: g3927863.
  atgaatgaaataagtcctacacgacctataaacattaatatcattaatacagccttggta
5 atgtcgttgacatcgcttgtcacacctaagcttaaaccacatgtaccaaacgcagacatt
  acttcaaagaaaatttgtaaaaaggacaacttgccttgttcagtggcagatataataatc
  atgctaataaatgtaattaatgacgccatagtaaatacggcaaatgatctttgtacatcc
  ataatatgtacttctctgttgaaaattttaatacctgttttgtcaccagtattattaaaa
  ttaataacaaacaaaattaaaatagcaaatgttgttgttctgattccacctcctacagaa
10 ctgggagatgatcctataaacatcaataatcccattacaatatttgttgcgtcgctgaaa
  tgtgacacatctatcgtttgcaaacctgcacttctggtcgttgatgattggaacaatgca
  taa

Sequence 3144
15 MNEISPTRPININIINTALVMSLTSLVTPKLKPHVPNADITSKKICKKDNLPCSVADIII
  MLINVINDAIVNTANDLCTSIICTSLLKILIPVLSPVLLKLITNKIKIANVVVLIPPPTE
  LGDDPININNPITIFVASLKCDTSIVCKPALLVVDDWNNA*

Sequence 3145
20 Contig_0660_pos_988_47
  is similar to (with p-value 3.0e-76)
  >sp:sp|Q46807|ARCL_ECOLI CARBAMATE KINASE-LIKE PROTEIN 1. >
  gp:gp|U28375|ECU28375_24 Escherichia coli K-12 genome; appro
  ximately 64 to 65 minutes. NID: g887800. >gp:gp|AE000370|AE0
25 00370_9 Escherichia coli K-12 MG1655 section 260 of 400 of t
  he complete genome. NID: g2367170.
  gtgagtgaaatggctaaaattgtagtagctttaggtggaaacgctttaggaaaatcacca
  caagaacaacttgaattagtaaaaaatacagctaaatccctagtaggattaattactaaa
  ggtcacgaaattgtgattagtcacggtaatggaccacaagtaggaagtattaaccttggt
30 ctgaattatgcagctgaacacgatcaaggtcctgcttttccatttgctgaatgtggcgct
  atgagtcaagcctacatcggctatcaacttcaagaaagtttacaaaatgaacttcattca
  atgggcatagataagcaagttgtcacactagttacccaagtagaagttgatgaaggcgat
  ccagcttttaatagtccaagtaaacccatcggtctgttctacactaaagaagaagcaaat
  cgtattcaacaggaaaaaggttatcaatttgtagaagatgctggtcgaggttaccgtcgc
35 gttgtaccatcaccacaaccaatatctattatcgaactggaaagtattaaaactctagta
  gaaaatgacacactcgtcatcgctgcaggtggaggtggtataccagtcattcgcgaacag
  catgatagctttaaaggtatagatgccgtcatcgataaagacaaaacaagtgcattatta
  ggtgctgatattcactgtgatcaactcattattttaacagcgattgattatgtttatatc
  aactatcatactgaccaacaacaagcacttaaaacaacaaatatagatacgcttaaaaca
40 tatattgaagaagaacaatttgccaaaggcagcatgctacctaaaatcgaatctgccatc
  tcctttattgaaaataatcctaacggtagcgtgctcatcacatcattaaatcaattagat
  gcagcactagaaggtaaaattggcacactcattacaaagtaa

Sequence 3146
45 VSEMAKIVVALGGNALGKSPQEQLELVKNTAKSLVGLITKGHEIVISHGNGPQVGSINLG
  LNYAAEHDQGPAFPPFAECGAMSQAYIGYQLQESLQNELHSMGIDKQVVTLVTQVEVDEGD
  PAFNSPSKPIGLFYTKEEANRIQQEKGYQFVEDAGRGYRRVVPSPQPISIIELESIKTLV
  ENDTLVIAAGGGGIPVIREQHDSFKGIDAVIDKDKTSALLGADIHCDQLIILTAIDYVYI
  NYHTDQQQALKTTNIDTLKTYIEEEQFAKGSMLPKIESAISFIENNPNGSVLITSLNQLD
50 AALEGKIGTLITK*

Sequence 3147
  Contig_0661_pos_4592_5317
  is similar to (with p-value 5.0e-26)
  >sp:sp|P23553|XYNC_CALSA ACETYL ESTERASE (EC 3.1.-.-). >pir
55 :pir|B37202|B37202 acetylesterase (EC 3.1.1.6) (XynC) - Cald
  ocellum saccharolyticum >gp:gp|AF005383|AF005383_9 Caldicell
  ulosiruptor saccharolyticus putative transport protein (XynG
  ), putative transport protein (XynH), xylanase (XynF), xylan

ase (XynE), xylanase (XynD), xylanase (XynA), acetylxylosida
se (XynC) and xylanase (XynB) genes, complete cds. NID: g264
5411. >gp:gp|M34459|CDCXYNAB_2 C.saccharolyticum xylanase A
(XynA), beta-xylosidase (XynB) and acetyl esterase (XynC) ge
5   nes, complete cds. NID: g144295.
atgcatcaaacgattaatgtcattcttccagaagataaaagttactttgatacaaatgaa
aatgcgaaaccattaaaaactatgttattgctacatggtttatcaagtgatacttcctct
tatatgagatatacgagtatagaacgctatgcgaatacccaccaactagcagtggtgatg
cctaatgctgatcatagtttctattcaaatatggcttatggacatagttattatgactat
10   atactagaagtttatgattatgttcatcaaatatttccattgtctaaaaagagagaagat
aattttatagcaggtcactctatgggaggttatggtgcaatcaaatttgcattaacgcaa
agttatcgtttctcaaaagccgctatgctttcagcgccatatgatgtttctatgattggt
caatatcaatggtatgattttactccagaagcgattgtaggtaatacgcaacatgtcgcg
gggacatcttttgatccatactatttagttgaacaagcaatagacaatggacaaacgtta
15   ccacaactatatattacttgtggaactgaagatgaattgtatcaaggtaatattgatttt
gtgaactatttagatgaaaaaggtatttcatatcaatttaaaaaagcgccaggtcatcac
gattatgcattttgggataaagcaatagaagatgtcattgaccgttttacatcatcacat
atttaa

20   Sequence 3148
MHQTINVILPEDKSYFDTNENAKPLKTMLLLHGLSSDTSSYMRYTSIERYANTHQLAVVM
PNADHSFYSNMAYGHSYYDYILEVYDYVHQIFPLSKKREDNFIAGHSMGGYGAIKFALTQ
SYRFSKAAMLSAPYDVSMIGQYQWYDFTPEAIVGNTQHVAGTSFDPYYLVEQAIDNGQTL
PQLYITCGTEDELYQGNIDFVNYLDEKGISYQFKKAPGHHDYAFWDKAIEDVIDRFTSSH
25   I*

Sequence 3149
Contig_0667_pos_3956_3216
is similar to (with p-value 2.0e-20)
30   >gp:gp|Y12813|BPA2INT_1 Bacteriophage A2 rep, xis and int g
enes. NID: g3005824.
atgcaagatttacaagtatttaattttgaagatttaccagtaagaaaaatagaagtagat
ggagaaccatatttttttaggtaaagacgtggcagaaatattaggttacacaagatctgat
aatgcaattagaaatcatgttgatgatgaagataagctgacgcaccaagttagtgcatca
35   ggtcaaaaacgaaacatggtaatcatcaacgaatctggtttatacagcttaatctttgac
gctgctaaacaaagtaaaaacgaaagtattagaaagaaagctaaacgttttaaacgttgg
gtaaccgaagatgttttaccttccattcgtaaaacaggtacttatcaagttcctgataat
ccaatggacgcattgcaacttatgttcgacgcacaaaaacaaaccaaagaagaaatagca
actgttaaagcagatgttattgatatcaaagaaatcaaaagctagatgcaggagaatac
40   ggattgataacaaaaacagttcatcaacgcgttgcttatatcagacaaattcacggacta
cctaataataaagaagttaacaaacctttatatagagatattaacagtaacgtaaatacg
atggctggtattaaaacaagaacacaattaaaacaaaaacatttcgatgacgtaatgaat
atgatcacaaattggtttccatctcaatcaacaatgtatgtcatcaaacaattagaaatg
gactttgaaaacgaagtataa
45

Sequence 3150
MQDLQVFNFEDLPVRKIEVDGEPYFLGKDVAEILGYTRSDNAIRNHVDDEDKLTHQVSAS
GQKRNMVIINESGLYSLIFDAAKQSKNESIRKKAKRFKRWVTEDVLPSIRKTGTYQVPDN
PMDALQLMFDAQKQTKEEIATVKADVIDIKENQKLDAGEYGLITKTVHQRVAYIRQIHGL
50   PNNKEVNKPLYRDINSNVNTMAGIKTRTQLKQKHFDDVMNMITNWFPSQSTMYVIKQLEM
DFENEV*

Sequence 3151
Contig_0670_pos_3378_4094
55   is similar to (with p-value 1.0e-45)
>gp:gp|AB014075|AB014075_8 Clostridium histolyticum genes f
or hypoxanthine-guanine phosphoribosyl-transferase (HGPRTase
), GTPase and 12 ORFs, complete and partial cds. NID: g38688
63.

```
atgggacgtaaatggaacaacattaaagagaaaaaagcccaaaaagataaaaatactagt
agaatatatgccaaatttggtaaagaaatatatgtagctgcaaagtctggtgagcctaat
ccagagtcaaatcaaactttaagattagtattagaacgtgcaaaaacatattcagtacct
aatcatattatagatagagctattgataaggctaaaggcgctggtgacgaaaactacgat
cacttaagatatgaaggttttggtccgaatggttcaatgcttatagttgacgcattaaca
aacaatgtaaatcgtacagcatcagatgtacgtgctgcgttcggtaagaatggaggaaat
atgggagtatctggttcagtagcttatatgtttgaccatactgcaacctttggtgtagaa
ggtaaatctgtagatgaagtcttagaaacactaatggagcaagatattgatgtaagagat
gtaattgatgacaatggcttgactattgtttacgcagaaccagatcaatttgcacaagtt
caggatgcattacgtgaagctggcgttgaggaatttaaagtagcagagtttgaaatgtta
cctcaaactgatattgagttgtctgaagaggatcaagctattttgaaaaattaatcgat
gcacttgaagacttggaagatgttcaaaatgttttccataatgtagatttaaaataa
```

Sequence 3152

MGRKWNNIKEKKAQKDKNTSRIYAKFGKEIYVAAKSGEPNPESNQTLRLVLERAKTYSVP
NHIIDRAIDKAKGAGDENYDHLRYEGFGPNGSMLIVDALTNNVNRTASDVRAAFGKNGGN
MGVSGSVAYMFDHTATFGVEGKSVDEVLETLMEQDIDVRDVIDDNGLTIVYAEPDQFAQV
QDALREAGVEEFKVAEFEMLPQTDIELSEEDQAIFEKLIDALEDLEDVQNVFHNVDLK*

Sequence 3153
Contig_0673_pos_2813_0
is similar to (with p-value 2.0e-34)
 >gp:gp|X81475|MHLMP_1 M.hominis lmp1 and lmp2 genes. NID: g
587470.

```
atgactgaagcaacaattcaaaattataacgctaaacgtcaaaaagcagagcaagttata
caaaatgcaaataaaattattgaaaacgctcaacctagtgtacaacaagtgtctgatgag
aaatctaaggtagagcaagcactcagtgaattgaacaacgccaaatcagcgcttagagct
gataaacaagaattacagcaagcatataatcagttgattcaaccaacggatttaaataat
aagaaaccagcttctatcactgcgtacaatcaaagatatcaacaatttagtaacgaattg
aacagcactaaaacaaatacagatcgcattttaaaagagcaaatccaagtgtagctgat
gtcaacaatgcactaaataaagtaagagaagtacaacaaaaattaaacgaagccagagca
cttttacaaaataaagaagataatagtgcactagttcgagccaaagaacaacttcaacag
gcagttgaccaagtcccttcaacagaaggtatgacgcaacaaactaaagatgattacaat
tcaaaacaacaagctgctcaacaagaaatatcaaaagcacaacaagttatcgataatggc
gatgcgactacacaacaaatttctaacgccaaaacaaatgttgaacgcgctttagaagca
ttaaataatgcaaaaactggtttaagagcagataaagaggaacttcaaaatgcatataat
caattaactcaaaatattgatacgagcggtaaaacgcctgcaagtatcaggaaatacaat
gaagctaagtcacgtattcaaactcaaattgattcagctaaaaataaagcaaacagtatt
ttaacaaatgacaatcctcaagtatcacaagtgactgctgcgttaaacaaaataaaagct
gttcaacctgaattagataaagcgatagcaatgcttaaaaataaagagaataataatgca
ttggttcaagcgaaacaacaacttcaacaaattgttaatgaagtagatccaacacaaggc
atgacaacagatactgctaataactataaatcaaaaaaacgtgaagctgaagatgaaata
caaaaagctcaacaaatcattaacaatggcgatgccactgagcaacaaattactaacgaa
acaaatagagtaaatcaagcgattaatgcaataaacaaagccaaaaacgatttacgtgct
gataagtctcaattggaaaatgcttataaccaattaatacaaaatgttgatacaaatggt
aaaaaacctgctagtattcaacaataccaagctgctcgacaagctattgagacgcaatac
aataacgctaaatcagaagcacatcaaattcttgaaaatagtaacccttcagttaatgaa
gtagcacaagcattacaaaaagttgaagctgtacaacttaaagttaatgacgcgattcat
atgcttcaaaataaagagaataatagtgcacttgtcacagctaaaaatcaacttcagcaa
gcagttaatgatcaaccattaacaacaggtatgactcaagattctattaataactatgta
gctaagagaaatgaggctcaaagtgctatcagaaatgcagaagctgtcatcaacaatggc
gatgcaactgcaaaacaaatttcagacgagaaatctaaagttgaacaagcactagcacat
ttgaatgatgctaaacagcaattaactgcagatactactgaattacaaacagcagttcaa
caattaaacagaagaggcgatacaaataataaaaagccaagaagtatcaatgcatataat
aaagcaattcaatcattagaaacacaaattacttctgctaaagataatgccaacgctgtg
atacaaaaacctatacgtactgttcaagaggtaaataatgcattacaacaagtaaatcag
ttgaatcaacaattaactgaagcaattaatcaacttcaaccgctatcaaataatgatgca
ttaaaagctgcaagattaaatttagaaaataaaattaatcaaactgtacaaactgatggt
atgacacaacaatctatagaggcttatcaaaacgctaaacgcgtagcccaaaatgaatct
```
```

```
        aacactgctttagcattaattaataacggcgatgccgatgaacaacaaattacaactgaa
        acagaccgagtcaatcagcaaactacaaacttaactcaagcaattaacgggttaacagtt
        aataaagaaccattagaaaccgctaaaacagcgttacaaaataacatcgaccaggtacct
        agtacagatggtatgactcagcaatctgttgcaaattataatcaaaaactacaaatagct
  5     aaaaacgaaattaacacaattaataacgttttagcgaacaatctagatgttaatgcaatc
        aaaacgaataaagcagaagcggaacgaatcagtaacgatttaacacaagctaagaataac
        ttacaagttgatactcaacctttagaaaaaataaaaagacaacttcaagatgaaattgat
        caaggtactaacacagatggaatgactcaagattcagtggataattacaatgatagctta
        agtgcagcaattatagaaaaaggcaaagtaaataaattacttaaacgtaatccgacagta
 10     gaacaagttaaagagagcgttgctaatgcacaacaagtcatacaagatttacaaaatgct
        cgaacttcacttgttccagacaaaactcaacttcaagaagctaaaaatagattagaaaac
        agtattaaccaacaaacagatactgacggcatgactcaagattcgcttaacaattataat
        gataaattagcaaaagctagacaaaaccttgaaaaaatatctaaagttttaggtggtcaa
        cctactgtagctgaaattagacaaaatacagatgaagcaaatgcacataaacaagcatta
 15     gacactgcacgttctcaacttacattaaatagagagccatatatcaatcatattaataat
        gaaagtcatttaaataacgcgcaaaaagataattttaaagctcaagttaactcagcacct
        aatcataatactttagaaacgattaaaaataaggctgatactttaaatcaatctatgaca
        gcattaagtggaagtattgcagattacgaaaatcaaaaacaacaagaaaattatttagat
        gcatctaacaataaacgtcaagactatgacaatcagtcaatgcggctaaaggtattttta
 20     aaccaaactcaaagtccgacaatgagtgctgatgtgattgatcaa
```

```
MTEATIQNYNAKRQKAEQVIQNANKIIENAQPSVQQVSDEKSKVEQALSELNNAKSALRA
DKQELQQAYNQLIQPTDLNNKKPASITAYNQRYQQFSNELNSTKTNTDRILKEQNPSVAD
VNNALNKVREVQQKLNEARALLQNKEDNSALVRAKEQLQQAVDQVPSTEGMTQQTKDDYN
SKQQAAQQEISKAQQVIDNGDATTQQISNAKTNVERALEALNNAKTGLRADKEELQNAYN
QLTQNIDTSGKTPASIRKYNEAKSRIQTQIDSAKNKANSILTNDNPQVSQVTAALNKIKA
VQPELDKAIAMLKNKENNNALVQAKQQLQQIVNEVDPTQGMTTDTANNYKSKKREAEDEI
QKAQQIINNGDATEQQITNETNRVNQAINAINKAKNDLRADKSQLENAYNQLIQNVDTNG
KKPASIQQYQAARQAIETQYNNAKSEAHQILENSNPSVNEVAQALQKVEAVQLKVNDAIH
MLQNKENNSALVTAKNQLQQAVNDQPLTTGMTQDSINNYVAKRNEAQSAIRNAEAVINNG
DATAKQISDEKSKVEQALAHLNDAKQQLTADTTELQTAVQQLNRRGDTNNKKPRSINAYN
KAIQSLETQITSAKDNANAVIQKPIRTVQEVNNALQQVNQLNQQLTEAINQLQPLSNNDA
LKAARLNLENKINQTVQTDGMTQQSIEAYQNAKRVAQNESNTALALINNGDADEQQITTE
TDRVNQQTTNLTQAINGLTVNKEPLETAKTALQNNIDQVPSTDGMTQQSVANYNQKLQIA
KNEINTINNVLANNLDVNAIKTNKAEAERISNDLTQAKNNLQVDTQPLEKIKRQLQDEID
QGTNTDGMTQDSVDNYNDSLSAAIIEKGKVNKLLKRNPTVEQVKESVANAQQVIQDLQNA
RTSLVPDKTQLQEAKNRLENSINQQTDTDGMTQDSLNNYNDKLAKARQNLEKISKVLGGQ
PTVAEIRQNTDEANAHKQALDTARSQLTLNREPYINHINNESHLNNAQKDNFKAQVNSAP
NHNTLETIKNKADTLNQSMTALSGSIADYENQKQQENYLDASNNKRQDYDNAVNAAKGIL
NQTQSPTMSADVIDQ
```

```
        atgtctaagttatcaattggacaaaatgatgttgtcaaaacaagtactaaaagacgcaat
 50     ttaatcttcaaagtcaatccgacttatcagcgacaaaaatttgttgtggattatgttgca
        aatcatgaaggacaggcaggaatcattttattgttccactcgtaagcaggtagaagaatta
        cacgaagctctaaatagtgaaaaaattaagagcacaatttatcatgctggtttaacgaat
        aaagagagaattgaggcgcaaaatgatttcttgtatgatcgtgtagaggttgtcattgcg
        acaaatgcatttggtatgggtattgataaatcaaatgtacgttatgtcattcactataac
 55     atgcctggagatttggaatcttactatcaggaagctggacgcgcgggacgtgatggttta
        aaaagtgagtgtatccttttgtttagtgaacgagataagggattacatgagtattttatt
        actgtatcccaagctgatgatgactataaagataaaatgggcgaaaaattaacgaaaatg
        attcaatataccaaaacgaaaaagtgtttagaagcgacaattgttcattattttgaaccc
        aatgaaaatttagaggaatgcaatcaatgtagtaattgtatacaggaaaataaaacgtat
```

788

```
    gatatgactcgtgaagcgaaaatgattattagctgtattgctcgaatgaagcaacaggaa
    aattatagtgttattatacaagtttttacgtggagaagtgacagattatataaaacaccat
    cattataatgaattaacgacacatgggttgatgaaaaattatacaacatctgagttatca
    cacttaattgatgagctacgtttcaaaggatatttaaatgaaaatgatgaaattcttatg
 5  tgtgatacatcagtgaaacaattactaaataatcataccaaggtttataccactccgttc
    aaacaaaaaactaaagagaaggtatttatcaacactgttgaaggtgtggatagagcgtta
    tatcgtgagcttgttgatgtacgtaaacagctaagtgataaacttggaatagcacctgta
    agtatattttctgattacacgctcgaagaatttgctaagcgtaaacctgaatcgaaacaa
    gaaatgattgctattgatggtgtaggtagttataaattaaagcattattgtcctaagttt
10  atcgaaaccatacaaagctataaaactagaatataa
```

Sequence 3156
MSKLSIGQNDVVKTSTKRRNLIFKVNPTYQRQKFVVDYVANHEGQAGIIYCSTRKQVEEL
HEALNSEKIKSTIYHAGLTNKERIEAQNDFLYDRVEVVIATNAFGMGIDKSNVRYVIHYN
MPGDLESYYQEAGRAGRDGLKSECILLFSERDKGLHEYFITVSQADDDYKDKMGEKLTKM
IQYTKTKKCLEATIVHYFEPNENLEECNQCSNCIQENKTYDMTREAKMIISCIARMKQQE
NYSVIIQVLRGEVTDYIKHHHYNELTTHGLMKNYTTSELSHLIDELRFKGYLNENDEILM
CDTSVKQLLNNHTKVYTTPFKQKTKEKVFINTVEGVDRALYRELVDVRKQLSDKLGIAPV
SIFSDYTLEEFAKRKPESKQEMIAIDGVGSYKLKHYCPKFIETIQSYKTRI*

Sequence 3157
Contig_0683_pos_4043_4831
is similar to (with p-value 1.0e-70)
  >sp:sp|P23355|PTFB_XANCP PTS SYSTEM, FRUCTOSE-SPECIFIC IIBC
  COMPONENT (EIIBC-FRU) (FRUCTOSE- PERMEASE IIBC COMPONENT) (
  PHOSPHOTRANSFERASE ENZYME II, BC COMPONENT) (EC 2.7.1.69) (E
  II-FRU). >pir:pir|B40944|B40944 phosphotransferase system en
  zyme II (EC 2.7.1.69), fructose-specific - Xanthomonas campe
  stris pv. campestris >gp:gp|M69242|XANFRUKAA_3 X.campestris
  1-phosphofructokinase (fruK) and PTS enzyme-II fructose (fru
  A) genes, complete cds. NID: g155366.

```
    gtggttttgttgagtaactcaaaaaaagtagccgttgtcacaggtgcagcacaaggtatt
    ggcttgaaaattgctgagcgtctatttgaagacggatatagcatcgcgcttgtagacttt
    aatgaagaggtagctaaagagtcagctgaaaaattatcaaaagaagggcaagaggcagtt
    gcttttaaagcagacgtttcaaatcgcgatcaagtatttagtgtgttaaatcaagtcgtt
    gaacactttggcgatttaaatgtcctagttaataatgctggtcttggaccaatgacacca
    attgaatcagtaacacctgaacaatttaatcaagttgtaggtgttaacgtaggtggtgta
    ttctggggtatccaagctgcaattgaacaatttgataaattaggacatggcggtaaaatt
    atcaatgccacatctcaagcaggtgttgaaggtaatgctggcttatctctatatagcagt
40  actaagtttgctgttagaggattgactcaagtagcagcacgtgatttagctgagaaaaat
    attacagtcaacgcattcgcacctggtattgttgaaacaccaatgatgaaaggtatcgct
    gaaaagcttgctgaggaaaataaccaaccaatggaatggggttggaaacaatttacagat
    caaattgccttaaaacgcttatctaaacctgaagatgtagctaatgtagtaagcttctta
    gcaggtagcgattcagattatattactggccaaacaatcatcgttgacggtggtatgaga
45  ttccactaa
```

Sequence 3158
VVLLSNSKKVAVVTGAAQGIGLKIAERLFEDGYSIALVDFNEEVAKESAEKLSKEGQEAV
AFKADVSNRDQVFSVLNQVVEHFGDLNVLVNNAGLGPMTPIESVTPEQFNQVVGVNVGGV
FWGIQAAIEQFDKLGHGGKIINATSQAGVEGNAGLSLYSSTKFAVRGLTQVAARDLAEKN
ITVNAFAPGIVETPMMKGIAEKLAEENNQPMEWGWKQFTDQIALKRLSKPEDVANVVSFL
AGSDSDYITGQTIIVDGGMRFH*

Sequence 3159
Contig_0683_pos_0_1268
is similar to (with p-value 3.0e-73)
  >sp:sp|Q48436|BUDC_KLEPN ACETOIN(DIACETYL) REDUCTASE (EC 1.
  1.1.5) (ACETOIN DEHYDROGENASE) (AR). >gp:gp|D86412|D86412_1
  Klebsiella pneumoniae gene for meso-2,3-butanediol dehydroge

nase (D-acetoin forming), complete cds. NID: g1468938.
gtgctcacttctaaagaaatcaaagaagctgatggaatcatcattgctgccgatagacag
gtagatttatcaaggtttaatggtaaacctctcatcaatgaaagtgtacgtgaaggtatt
catagacccaaggaattaatacaacgtgtgattgaccaagatgcacaaatttaccatgat
caaaatatttcttcaaatatgtctagagaccaggaagaatctcataaaagtaatattcaa
atggtatatcagcatttaatgaatggtgtttccttcatggttccatttatcgtcgttggc
ggtttactcatagctattgccttaactcttggaggacacaccactccaaaaggattagtt
atccccgaagattcattttggaaatctattgaaaatattggtagtttatcgtttaaattc
atggttcccatccttgctggttatatcgcggtgagtattgctgataagcctggtcttgtt
ccaggtatgattggtggtgccattgctgctgatggtagtttatatggaagtgaagcagga
gccggtttccttggtggtatcgtcgcaggtttcttagcgggctatattgcaaaatggatt
aaacagattaaagttcctaaagctatggctcctattatgcctattattattatacctatt
ctatcttctttaatagttggtctcattttttatatttgtaataggcgcaccaatttcaaat
atatttggtgcattaacatcatggttaaaaggaatgcaaggtgctaacatcattattctt
gctcttattattggcgcgatgattgcttttgatatgggaggtccagtaaacaaagtagca
ttcttattcggttctgcattaattgctgaaggcaactacactgtgatgggaatggttgct
gtagcagtatgtacaccaccgattggtttaggtttagctacattgttcgtaaacaccaa
ttcaataaagcagaacaagaaatgggtaaggcatcatttacgatgggattatttggtatt
actgaaggggcaatccctttttgctgcacaagatcctctaagaatcattccagccaatatg
attggcgcgatgattgcttcagtaatagcggcggttggaggtgtcggtgataaagttgct
catggaggtcctattgtcgctgtactaggtggaataagtaatattttatggttctttata
gctgttgtcgttggaagcttagtaaccatgttcacagtcttgttatttaagcgtcacacc
cctgcTTT

Sequence 3160
VLTSKEIKEADGIIIAADRQVDLSRFNGKPLINESVREGIHRPKELIQRVIDQDAQIYHD
QNISSNMSRDQEESHKSNIQMVYQHLMNGVSFMVPFIVVGGLLIAIALTLGGHTTPKGLV
IPEDSFWKSIENIGSLSFKFMVPILAGYIAVSIADKPGLVPGMIGGAIAADGSLYGSEAG
AGFLGGIVAGFLAGYIAKWIKQIKVPKAMAPIMPIIIIPILSSLIVGLIFIFVIGAPISN
IFGALTSWLKGMQGANIIILALIIGAMIAFDMGGPVNKVAFLFGSALIAEGNYTVMGMVA
VAVCTPPIGLGLATFVRKHQFNKAEQEMGKASFTMGLFGITEGAIPFAAQDPLRIIPANM
IGAMIASVIAAVGGVGDKVAHGGPIVAVLGGISNILWFFIAVVVGSLVTMFTVLLFKRHT
PAX

Sequence 3161
Contig_0687_pos_5614_4433
is similar to (with p-value 4.0e-55)
  >sp:sp|Q44681|RISB_BACAM 6,7-DIMETHYL-8-RIBITYLLUMAZINE SYN
THASE (EC 2.5.1.9) (DMRL SYNTHASE) (LUMAZINE SYNTHASE) (RIBO
FLAVIN SYNTHASE BETA CHAIN). >gp:gp|X95955|BARIBGENS_4 B.amy
loliquefaciens ribB, ribG, ribA, ribH & ribT genes. NID: g15
92687.
atgcaattcgatacaattgagttggctatagaggctttaagaaatggagagagcattatt
gtagttgacgatgaagatagagaaaatgaaggagatcttgtagctgttacggaatggatg
gatgataataccattaatttttatggctaaagagggtcgtggtctgatttgtgcaccaatt
gataaatctatagctgaaagattaaaactacaatctatggagcaaaataacactgatatt
tatggcacacatttttactgtaagcattgatcattataaaactactacaggaatcagtgca
catgaacgtacacaaacggctagagcactcatagatgaaaatactaatcctgaagatttt
catcgtccggggcacttatttccacttatagcaaaagagaatggtgtgttaacacgtaat
ggtcatactgaagctgccgtagatttggcacggttaacaggagcacaaccagctggagta
atctgcgaaattatgaatgatgatgggacaatggctaagggtgaagatctccagtcattt
aaagaacgccaccatttaaaaatgattactataaaaagtttggttgcttttcgtaaggct
gttgaacttaatgttaatcttaaggcaaaggtcaagatgccaactgattttggtcatttt
gatatgtatggatttacaacggattatagcgatgaagaaatcgtagctattgttaaagga
gatttaaaaagcaatcctaatgtacgtatgcattctgcttgtctgactggggatatttttt
catagtcaaagatgtgattgcggggcacaacttgaagcgtcaatgaaatatattgacgaa
catggtggaatgattatttatttacctcaagaaggtagaggaatagggttaattaataag
ttgcgcgcatatgagttgatagaaaaaggttatgatacagttactgcaaatcttgctctt
ggttttgatgaggatttgagagattatcatgttgcagctgaaatattaaagtattttgat

```
ataagtgaaattaacttgctcagcaataatcctaaaaaatttgaaggtttagaagattac
ggcattgagatcgtagatagaattgaacttatcgttccagaaacacaatataaccatagt
tatatggaaactaaaaaaaataaaatgggacatttaatatag
```

5 Sequence 3162
MQFDTIELAIEALRNGESIIVVDDEDRENEGDLVAVTEWMDDNTINFMAKEGRGLICAPI
DKSIAERLKLQSMEQNNTDIYGTHFTVSIDHYKTTTGISAHERTQTARALIDENTNPEDF
HRPGHLFPLIAKENGVLTRNGHTEAAVDLARLTGAQPAGVICEIMNDDGTMAKGEDLQSF
KERHHLKMITIKSLVAFRKAVELNVNLKAKVKMPTDFGHFDMYGFTTDYSDEEIVAIVKG
10 DLKSNPNVRMHSACLTGDIFHSQRCDCGAQLEASMKYIDEHGGMIIYLPQEGRGIGLINK
LRAYELIEKGYDTVTANLALGFDEDLRDYHVAAEILKYFDISEINLLSNNPKKFEGLEDY
GIEIVDRIELIVPETQYNHSYMETKKNKMGHLI*

Sequence 3163
15 Contig_0687_pos_4420_3959
  >sp:sp|P51695|GCH2_BACAM GTP CYCLOHYDROLASE II (EC 3.5.4.25
) / 3,4-DIHYDROXY-2-BUTANONE 4- PHOSPHATE SYNTHASE (DHBP SYN
  THASE). >gp:gp|X95955|BARIBGENS_3 B.amyloliquefaciens ribB,
  ribG, ribA, ribH & ribT genes. NID: g1592687.
20 atgaattttgaaggtaaattagttggtaaggatttaaaaaattgcgattgttgttagtaga
tttaatgattttattactacacgtctacttgaaggggctaaagatacacttattcgtcat
gaagtagaagatacaaatattgatgtagcttatgtgcctggcgcattcgaaattccactc
gttgcaaaaaaattagctcaaaaaggtgaatatgatgctgtgattacattaggatgtgtg
attagaggcgcaacttcacattatgactatgtatgtaatgaagtagctaaaggtgtttct
25 aaagcaaacgacatttcagatactccggtgattttttggagttctaacaactgaaagtatt
gaacaagcagttgaaagagctggtactaaagctggaaataaaggttcagaagcagcagtt
agtgcaatcgaaatggctaatctaattaagcaaatcaattaa

Sequence 3164
30 MNFEGKLVGKDLKIAIVVSRFNDFITTRLLEGAKDTLIRHEVEDTNIDVAYVPGAFEIPL
VAKKLAQKGEYDAVITLGCVIRGATSHYDYVCNEVAKGVSKANDISDTPVIFGVLTTESI
EQAVERAGTKAGNKGSEAAVSAIEMANLIKQIN*

Sequence 3165
35 Contig_0691_pos_489_154
  is similar to (with p-value 2.0e-21)
  >pir:pir|A55345|A55345 diamine N-acetyltransferase (EC 2.3.
  1.57) - Escherichia coli >gp:gp|D25276|ECOSN1A_1 Escherichia
  coli gene for spermidine acetyltransferase, complete cds. N
40 ID: g517104. >gp:gp|AE000254|AE000254_5 Escherichia coli K-1
  2 MG1655 section 144 of 400 of the complete genome. NID: g17
  87862.
  gtggtagagttattagaaattaactttatacatagaacttgtgaagtgttaattattatc
  gatccgcagtatgcaaataatgggtacgcgaaaaaagcctttaaaatggctattgactat
45 gcttttttagtattaaatatgaataaggtatacttatatgtggatattaagaatgagaaa
gcagtacatatctatcaaagtaataatttcgaaatagaaggaacgttaaaggaacacttc
tatacaaggggagaatatagagattgctatgtaatgggcttgttaaaaaggaattgggtt
aataagaatgatgatgatttgtctcatataagatga

50 Sequence 3166
VVELLEINFIHRTCEVLIIIDPQYANNGYAKKAFKMAIDYAFLVLNMNKVYLYVDIKNEK
AVHIYQSNNFEIEGTLKEHFYTRGEYRDCYVMGLLKRNWVNKNDDDLSHIR*

Sequence 3167
55 Contig_0692_pos_844_1881
  is similar to (with p-value 9.0e-27)
  >sp:sp|P45578|YGAG_ECOLI 19.3 KD PROTEIN IN EMRB-GSHA INTER
  GENIC REGION.
  atgcataataaacaaaagatattagatttatagaaaataataaatatgattatgttgaa
```

```
      ataagtcatcgtattcatgaacgccctgaattaggcaatgaagaaatttttgcatcgaga
      acattaattgaccaattaagagcaaatcgattcgaaatcgaacggatattgcaggacat
      gcaacaggatttatagcaacgtatgattctgatatgactggaccggttataggatttcta
      gctgaatatgatgctttacctggtcttggtcacgcatgcgggcataatattattggtact
  5   gctagcgtacttgctgcagtagcactaaaagaagtcgtcgatgaaattggtggtaaagta
      gtcgttttgggatgtcctgctgaagaaggtggggaaaatggctccgcaaaagcttcttat
      gttaaagcaggtgtcattgatgaaattgatgtagcattgatgattcatcctggaaatgaa
      acttatcgtacaattaatactttagctgtggatgttctagatattaaattctatggacgt
      agtgcgcatgcatctgaaaatgcagatgaagcattaaacgctttagatgcaatgatttca
 10   tatattaatggtatagcacagttaaggcaacacattaaaaaaggacaacgagttcacggg
      gttattttagacggtggtaaagcggctaatattatacctgattttacacatgcgagattt
      tacactcgagctacttcgcggagagaacttgatgtttttaactgaaaaagtaaaccaatt
      gcaagaggtgctgctattcaaactgggtgtgattttgaatttggtcctatccagaatggt
      gtaaacgaatttatcaaagcacctaaacttgatgatttatttgaaaaatatgcaactgaa
 15   ttaggagaagaagtgattgatgatgattttggctatggatctacagatacaggtaatgta
      agtcatgttgtaccaacaatacatccacatattaaaattggttctcgaaatcttgttagg
      acatacccaccgctttag
```

      Sequence 3168
 20   MHNKQKILDFIENNKYDYVEISHRIHERPELGNEEIFASRTLIDQLRANRFEIETDIAGH
      ATGFIATYDSDMTGPVIGFLAEYDALPGLGHACGHNIIGTASVLAAVALKEVVDEIGGKV
      VVLGCPAEEGGENGSAKASYVKAGVIDEIDVALMIHPGNETYRTINTLAVDVLDIKFYGR
      SAHASENADEALNALDAMISYINGIAQLRQHIKKGQRVHGVILDGGKAANIIPDFTHARF
      YTRATSRRELDVLTEKVNQIARGAAIQTGCDFEFGPIQNGVNEFIKAPKLDDLFEKYATE
 25   LGEEVIDDDFGYGSTDTGNVSHVVPTIHPHIKIGSRNLVRTYPPL*

      Sequence 3169
      Contig_0692_pos_609_148
      is similar to (with p-value 8.0e-17)
 30   >gp:gp|AF006687|AF006687_1 Enterobacter agglomerans indole-
      3-acetyl-L-aspartic acid hydrolase gene, complete cds. NID:
      g2654566.
      atgaatgtagaaagctttaatttagaccatactaaggttgttgcaccttttattcgtcta
      gccgggactatggaaggtcttaatggtgatgtcatacacaaatatgacattcgtttcaaa
 35   cagcccaataaggaacatatggatatgcctggtctacattccttagagcatttaatggca
      gaaaacattagaaatcatactgataaagtagtagatttaagtcctatgggttgtcaaact
      ggattctatgtttcatttattaatcatgacgactacgatgacgtattaaatattatcgat
      caaacattgcatgatgtgttaaatgctagcgaagtcccagcttgtaatgaggttcaatgt
      ggttgggctgcaagtcattctttagaaggtgctaaaacaattgctcaagcatttttagat
 40   aaagagagcaatggaatgacatctacggagaaggtaaataa

      Sequence 3170
      MNVESFNLDHTKVVAPFIRLAGTMEGLNGDVIHKYDIRFKQPNKEHMDMPGLHSLEHLMA
      ENIRNHTDKVVDLSPMGCQTGFYVSFINHDDYDDVLNIIDQTLHDVLNASEVPACNEVQC
 45   GWAASHSLEGAKTIAQAFLDKREQWNDIYGEGK*

      Sequence 3171
      Contig_0693_pos_5566_6633
      is similar to (with p-value 9.0e-84)
 50   >gp:gp|X94433|BCASPAMIN_1 B.circulans aspartate aminotransf
      erase gene. NID: g1147556.
      atgctaatagataaagcaagatcatttattcagaccatgtatagcgaattaaaatataat
      actaatgaaattgaaaatagaatgaaagagattgagcaagaaattaacttgactggtagt
      tacacacatacttatgaagaattatcttacggtgcaaaaatggcatggagaaactcaaat
 55   cgttgtattggtagactgttttggaattctttaaatgttaaagatgcccgagatgtatgt
      gacgaaaaagaatttataaaatttatacatacacatattaaagaagctactaacggcgga
      aaaatcaaaccatatattacaatttttagtcctgaagatacacctaaaatttataataat
      cagttgattcgttatgctggttatgaaaatgttggcgatccatctgaaaaaaaggttact
      cgtttagctgaacatctaggttggaaaggtaaaggttcaaattttgatattttacctctg

```
atttatcaattgcctaacgacactataaaaatacacgaacttccaaatgatattgtaaaa
gaagtttctatacatcatgaacactatcccaagctttcaaaattaggtttaaaatggtat
gcggtacctattatttcaaatatggatttaaaaatcggtggtattacttaccctacagca
ccttttaatggatggtatatggtaaccgaaattgctgtacgtaatttcacagacacctat
```
5
```
cgttataatcttttagaaaaagttgcagaagcttttgaatttgatacacttaaaaataat
tcatttaataaagatcgagcactcgtagagttaaatcatgctgtgtatcattcatttaaa
gctgatggtgtttctattgttgaccacttaactgcagcgaagcaatttgaaatgtttgaa
cgaaatgaacatcaacaaacagaaatgttactggtaagtggtcttggctggcacctcca
ctttcaccaactttaacttctaactatcatcatggatatgataatacaatgcatcatacg
```
10
```
aatttcttctataaaaaagaagaacctatgaagtgccctttccattaa
```

Sequence 3172
MLIDKARSFIQTMYSELKYNTNEIENRMKEIEQEINLTGSYTHTYEELSYGAKMAWRNSN
RCIGRLFWNSLNVKDARDVCDEKEFIKFIHTHIKEATNGGKIKPYITIFSPEDTPKIYNN
15 QLIRYAGYENVGDPSEKKVTRLAEHLGWKGKGSNFDILPLIYQLPNDTIKIHELPNDIVK
EVSIHHEHYPKLSKLGLKWYAVPIISNMDLKIGGITYPTAPFNGWYMVTEIAVRNFTDTY
RYNLLEKVAEAFEFDTLKNNSFNKDRALVELNHAVYHSFKADGVSIVDHLTAAKQFEMFE
RNEHQQNRNVTGKWSWLAPPLSPTLTSNYHHGYDNTMHHTNFFYKKEEPMKCPFH*

20 Sequence 3173
Contig_0693_pos_12681_13970
>gp:gp|U13618|SEU13618_2 Staphylococcus epidermidis 9759 he
at shock protein 10 (hsp10) and heat shock protein 60 (hsp60
) genes, complete cds. NID: g535340.
25
```
atgatgaatccattagcccaaaaattgaatgatgaaataaagcaatcaagtccagaagta
ttagatatgatgtcacaattaggtaaagatatgttttatccaaaaggaatttttatcgcaa
tctgccgaagcaaaacgcacaacatataatgctactattggtatggcaaccaaaaaagaa
ggtaaaatgtacgcaaattcacttaaccaaatgtttaatgaccttacaccggatgaaatt
ttcccatatgcacctcctcaaggtgtagaggaattacgtgatttatggcagaaaaaaatg
```
30
```
cttaaagaaaatcccgacttaaagtctaaatctatctctcgtcccatcgttacaaatgct
ctcacgcacggtctttctctagtagctgatttatttgtagatacagatgataccgtctta
ttaccgacacacaactggggtaattataaacttgtatttagcacacgtcacggtgctcat
atcaatacgtattctattttttgatgactcaggtcacttcactacatctgaacttgtaaaa
acattaaaagaatataaaaaagacaaagtgattattattttaaattatcctaataaccca
```
35
```
actggttacacaccaaataaaaaaagaagttaatactattgtaaatgcaattgaagaacta
gctaataaaggtactaaagtagtaactgttgtcgatgatgcatactatgggttattttat
gaagaagtttaccaacagtcaattttcacggctttaacacaggtgaaatcttctaacctt
ttaccagtgcgtttggatggagctactaaagaattcttctcttggggggttccgagttggc
tttatgacgtttggaattgatcatgaaacgttaaaaaatgcgctagaagctaaagtaaaa
```
40
```
ggattaattcgtagcaatatttcaagttctccactaccttctcaaagtgcaatcaaacat
gtacttaaatatcatgagcaatttgataaagaaatcgatcaaaatatcaatattttaaaa
gaacgctacgaagtaactaaacaagtagtgtatgataataaatatgccaaatattggcaa
gcctatgactttaattcaggatactttatgtcattgaaattaaatcaagtcgatccagaa
gaattacgtaaacatttaattaataactattcaattggtattattgctttaaatagtaca
```
45
```
gatattcgtattgcctttagttgtgtagaaaagaagatattccttatgtctttgagtct
attgctaatgcaattgatgatattaaataa
```

Sequence 3174
MMNPLAQKLNDEIKQSSPEVLDMMSQLGKDMFYPKGILSQSAEAKRTTYNATIGMATKKE
50 GKMYANSLNQMFNDLTPDEIFPYAPPQGVEELRDLWQKKMLKENPDLKSKSISRPIVTNA
LTHGLSLVADLFVDTDDTVLLPTHNWGNYKLVFSTRHGAHINTYSIFDDSGHFTTSELVK
TLKEYKKDKVIIILNYPNNPTGYTPNKKEVNTIVNAIEELANKGTKVVTVVDDAYYGLFY
EEVYQQSIFTALTQVKSSNLLPVRLDGATKEFFSWGFRVGFMTFGIDHETLKNALEAKVK
GLIRSNISSSPLPSQSAIKHVLKYHEQFDKEIDQNINILKERYEVTKQVVYDNKYAKYWQ
55 AYDFNSGYFMSLKLNQVDPEELRKHLINNYSIGIIALNSTDIRIAFSCVEKEDIPYVFES
IANAIDDIK*

Sequence 3175
Contig_0693_pos_16689_15784

unknown; although very good hit (p=0) to B. subtilis genome

```
gtgcaggctagtcgaccaatttttaattgttgcggatgaagtagaaggcgatgcacttact
aatattgttttaaaccgtatgcgtggaacatttactgctgtagcagttaaagccccagga
```
5
```
tttggtgatcgacgtaaagcaatgttagaagacctagcaatattaactggtgctcaagtc
attactgatgatttaggtttagaacttaaagatgcatctcttgatatgctaggtactgct
aataaagttgaagtgactaaagatcatacaacagtcgtagatggtaatggtgatgaaat
aatattgatgctcgtgtaggtcaaattaaagcacaaattgaagaaactgattcagagttt
gataaagaaaaattacaggaacgtttggcaaaactagctggcggcgtagctgttatcaaa
```
10
```
gtaggggctgcaagtgaaacagagcttaaagaacgtaaattaagaattgaagacgcatta
aattcaacacgtgcggcggtggaagaaggtatcgttgctggtggtggtactgcgttagtc
aatatatatcaaaaagtaagtgaaattaaagcagaaggtgatgttgaaacgggtgttaat
atcgtattaaaagcattacaagcacctgttagacaaattgctgaaaatgcaggattagag
ggttcaattattgttgaacgtttaaaacatgctgaagcgggcgttggtttcaatgcagca
```
15
```
acaaatgaatgggttaatatgttagaagaaggtatagtagatccaactaaagtaactcgt
tcagcgttacaacatgcagcaagtgtagctgctatgttcttaacaactgaagcagtcgtt
gctagtattccagagccagaaaataatgaacaacctggaatgggtggcatgccaggtatg
atgtaa
```

20   Sequence 3176
VQASRPILIVADEVEGDALTNIVLNRMRGTFTAVAVKAPGFGDRRKAMLEDLAILTGAQV
ITDDLGLELKDASLDMLGTANKVEVTKDHTTVVDGNGDENNIDARVGQIKAQIEETDSEF
DKEKLQERLAKLAGGVAVIKVGAASETELKERKLRIEDALNSTRAAVEEGIVAGGGTALV
NIYQKVSEIKAEGDVETGVNIVLKALQAPVRQIAENAGLEGSIIVERLKHAEAGVGFNAA
25   TNEWVNMLEEGIVDPTKVTRSALQHAASVAAMFLTTEAVVASIPEPENNEQPGMGGMPGM
M*

Sequence 3177
Contig_0693_pos_5374_3905
30   is similar to (with p-value 3.0e-77)
>gp:gp|U59924|SSU59924_1 Sus scrofa nitric oxide synthase (
NOS) mRNA, complete cds. NID: g1762433.
```
gtgtaccaatataacgacgatagcttaatgttacacaatgatttatatcaaattaatatg
gctgaaagctactggaatgatggtatccatgaaagaatagcagtgtttgatttgtatttt
```
35
```
cgaaaaatgccatttaatagtggatatgcggtattcaacggattgaaacgcgttgtgaat
ttcatcgaaaactttgggtttacaaatgaagatatcacatatttaaaatcgataggttat
gaagaagattttctaaattacctaaaagatttgaaatttacagggaatattaaatctatg
caagaaggtgaaatttgtttttggtaatgagccattattaagagttgaagcacctttaatc
caagcgcaacttattgaaactattttgttaaatatcattaatttccaaacattaattgca
```
40
```
actaaagccagccgaattcgtcaaatagcaactcatgacactttgatggaatttggtaca
agaagagctcaagagatcgatgctgcactgtggggcgctagagcagcctttattggaggg
tttgattctacaagtaatgttagagcaggaaaactttttaatatacctgtatctggcaca
catgcacacgcactagtacaaacatatggtgatgagtatatagcattcaaaaagtatgct
gagcgacataaaaattgtgtgttcttagttgatacttttcatactttaaaatcaggagta
```
45
```
ccaaccgcaattaaggttgcaaaagagttaggagatactattaattttataggtatcaga
ttagattctggtgatattgcgtacctatctaaagaagctcgtagaatgttagatgaggct
ggttttacagaagctaaaattatcgcatcaaatgatttggatgagcagactattacaagt
ttaaaagcacaaggcgctaaagttgacggatggggagtaggtacaaaactgattacagga
tatgatcaaccagccttaggtgcagtttataaattggtttctattgaaacagatgatggc
```
50
```
acaatgagtgatcgcattaaattatcaaataatgctgagaaagttactacaccaggcaaa
aaaaatgtttatcgtattattaataataaaacaggcaaggctgagggcgactatattacg
ctagaaggtgaaaatcctaatgacgaatctccattgaaaatgttccatcctgttcacact
tacaaaatgaagtttattaaatcatttaaagcggttaatctacatcaatctatatttgaa
aatggcaaacttgtataccatcttccagatgaatatgaagctcaggactatcttaaaaat
```
55
```
aatttaagtatttttatgggaagaaaataaacgatatcttaacccgcaagattatccagta
gatttaagcactaaatgttgggaaaataagcataagcgtattttttgaagttgctgaacac
gttaaagagatggaggatgaaaatgagtag
```

Sequence 3178

VYQYNDDSLMLHNDLYQINMAESYWNDGIHERIAVFDLYFRKMPFNSGYAVFNGLKRVVN
FIENFGFTNEDITYLKSIGYEEDFLNYLKDLKFTGNIKSMQEGEICFGNEPLLRVEAPLI
QAQLIETILLNIINFQTLIATKASRIRQIATHDTLMEFGTRRAQEIDAALWGARAAFIGG
FDSTSNVRAGKLFNIPVSGTHAHALVQTYGDEYIAFKKYAERHKNCVFLVDTFHTLKSGV
5   PTAIKVAKELGDTINFIGIRLDSGDIAYLSKEARRMLDEAGFTEAKIIASNDLDEQTITS
LKAQGAKVDGWGVGTKLITGYDQPALGAVYKLVSIETDDGTMSDRIKLSNNAEKVTTPGK
KNVYRIINNKTGKAEGDYITLEGENPNDESPLKMFHPVHTYKMKFIKSFKAVNLHQSIFE
NGKLVYHLPDEYEAQDYLKNNLSILWEENKRYLNPQDYPVDLSTKCWENKHKRIFEVAEH
VKEMEDENE*

10

Sequence 3179
Contig_0694_pos_3485_4648
is similar to (with p-value 2.0e-36)
    >sp:sp|P31448|YIDK_ECOLI HYPOTHETICAL 62.1 KD PROTEIN IN EM
15   RD-GLVG INTERGENIC REGION. >gp:gp|L10328|ECOUW82_44 E. coli;
    the region from 81.5 to 84.5 minutes. NID: g290484. >gp:gp|
    AE000445|AE000445_8 Escherichia coli K-12 MG1655 section 335
    of 400 of the complete genome. NID: g1790105.
    gtgttcaataaaatgtttaaagttgatgaatacttaggtgtaagtagttcaactgctgtc
20   attattatttcatctattattggtataattggcattatttacttatttataggtggttta
    tcgttaagtgcttttagtgattcgatttatggcatggctcttaattatagctggacttgcg
    attacaatattaggtctaggtcaattaggagatggcaacttcctacatggtttcgacaaa
    atcgtgcaagacacgcctgagaaattgaatggtttggtaaggtggactcggatgttgta
    ccttggccaacccctattcttcggtatgttctttaacaatttattcttctggtgcgcaaac
25   cagatgatagttcaaaaagcactcgcagctaaaaatttaaaagaatctcaaaaaggtgca
    atatatttaagcttatttaaagtgttcggacctttatttacagtcttaccaggcgtagta
    gcatttaactattttaatggtagtcttgataaatcagataacgcttaccctgcacttgta
    acttcagtattaccagaatgggcattcggcttatttggtgcggttattttggtgcaata
    ttgagctcatttgttggctcattgaatagtacaactacactattaacactcgatttctat
30   aaacctattttttggaaaaaataaatcagataaacatattgctcgagtaggccatattgct
    actgtagtcattggagttattgttgtagcacttgcaccagtcatctcattattccctagt
    ggtctttatgcagtagttcaacagtttaatggtgtgtatagtatgccagtgctagcttta
    attttaatcgctttcttttctaaacgcacatctaaattgggcgctaaagtgacactcttc
    acacatataattttatacgctataatcagctttgtatttacagaaattaattatctatac
35   acatttagtgtattattctttgtagatttaattattattttgatttttaacaaagttaaa
    ctatctagtgagtttgatttaagcacgcaccaaccgaaagtagatatgacgccatggaaa
    tatcgttacgttgcaggtattattgttcttgctttagtagtagtaagttatattatcttc
    tcaccactcgtgttggcaaaataa

40   Sequence 3180
VFNKMFKVDEYLGVSSSTAVIIISSIIGIIGIIYLFIGGLSLSAFSDSIYGMALIIAGLA
ITILGLGQLGDGNFLHGFDKIVQDTPEKLNGFGKVDSDVVPWPTLFFGMFFNNLFFWCAN
QMIVQKALAAKNLKESQKGAIYLSLFKVFGPLFTVLPGVVAFNYFNGSLDKSDNAYPALV
TSVLPEWAFGLFGAVIFGAILSSFVGSLNSTTTLLTLDFYKPIFGKNKSDKHIARVGHIA
45   TVVIGVIVVALAPVISLFPSGLYAVVQQFNGVYSMPVLALILIAFFSKRTSKLGAKVTLF
THIILYAIISFVFTEINYLYTFSVLFFVDLIIILIFNKVKLSSEFDLSTHQPKVDMTPWK
YRYVAGIIVLALVVVSYIIFSPLVLAK*

Sequence 3181
50   Contig_0699_pos_1788_1444
is similar to (with p-value 4.0e-59)
    >pir:pir|I67760|I67760 transposase (insertion sequence IS10
) - Escherichia coli >gp:gp|S67119|S67119_2 BST=somatotropin
    ...BST/beta-Gal fusion protein [Escherichia coli, LBB84, pla
55   smid pXT107, IS10L/R-1, PlasmidTransposonInsertionMutant, 3
    genes, 1679 nt]. NID: g455674.
    atgcagattgaagaaaccttccgagacttgaaaagtcctgcctacggactaggcctacgc
    catagccgaacgagcagctcagagcgttttgatatcatgctgctaatcgccctgatgctt
    caactaacatgttggcttgcgggcgttcatgctcagaaacaaggttgggacaagcacttc

caggctaacacagtcagaaatcgaaacgtactctcaacagttcgcttaggcatggaagtt
ttgcggcattctggctacacaataacaagggaagactcactcgtggctgcaaccctgctt
actcaaaatctattcacacatggttacgttttggggaaattatga

5   Sequence 3182
MQIEETFRDLKSPAYGLGLRHSRTSSSERFDIMLLIALMLQLTCWLAGVHAQKQGWDKHF
QANTVRNRNVLSTVRLGMEVLRHSGYTITREDSLVAATLLTQNLFTHGYVLGKL*

     Sequence 3183
10   Contig_0702_pos_12730_11027
      >sp:sp|P16659|SYP_ECOLI PROLYL-TRNA SYNTHETASE (EC 6.1.1.15
) (PROLINE--TRNA LIGASE) (PRORS) (GLOBAL RNA SYNTHESIS FACTO
     R).
     atgaaacaatctaaagtatttataccaacgatgagagaagtccctgcagaggcggaagca
15   ttaagccatcgtttattattaaaagcagggttaattaaacagagtacaagtggtatatat
     agttacttaccacttgctacacgtgtactaaataatatatctaaatcatacgtgaagaa
     atggaaagtatagatgctgtagaaattcttatgccagctttacaacaagcagaattatgg
     gaagagtcaggacgttggagtgcatatggtccagaactaatgcgtttaaaagacagaaac
     ggacgtgaatttgcattaggacctactcatgaggaagtagtcacttctttagtaagagat
20   gaattaaaatcatataaacagttaccccttactttatttcaaattcaatccaaatataga
     gatgagaaaagaccacgctttggattattaagaggacgcgaatttcttatgaaagatgca
     tattccttccattcagatgaagcttcattagatgcaacttatcaagatatgtatcaagca
     tatagtcgcatattcaaacgtgtaggcatcaatgcaagaccggttgtggcagattcaggt
     gcaatagggggaagtcatacacacgagtttatggcattgagtgaaattggggaagataca
25   atagtttatagtaatgagagtgactatgcagcgaatattgaaaaggctgaagttgtttat
     catccttctcataagcattctgcacttgcggaattgactaaagttgagacgcccaatgtt
     aaaacagctcaagaagttgcagaatatttaaagagaccattagatgaaattgtaaaaact
     atgatctttaaaatagatggcgaatttattatgtttctagttcgtggacatcatgaatta
     aatgaagtgaaattaaaatcatatttcggcacggaacatgttgaaatggctactccagat
30   gaaattgttaatcttgtagatgccaatccggggtctcttggtcctatttttgataaagat
     attaaaatttatgccgataattacttacaagatttaaataactttgttgtaggagctaat
     gaagatcattatcactatataaatgtcaatattggtagagactttgatgtaacagaatac
     ggtgactttagattcattacacaaggtgagatgttaagtgatggctcgggagtagcacaa
     tttgctgaaggcattgaagtaggacaagttttcaaattagggacaaaatattctgaatca
35   atgaatgcaactttttctagataatcagggaaaagctcaaccactcattatgggctgttat
     ggtattggagtatcaagaacattaagtgcaattgttgaacaaaacaatgacgagaatgga
     attatttggccaaaatcagtaacgcctttcgatatccatctaattactatcaatcctaaa
     aaggatgatcaacgtactttaggtgatcaactttatcaaaaattaatggattcatacgat
     gttttatatgatgaccgaaaagaacgtgctggtgttaaatttaatgattcagacctaatt
40   gggttaccggtacgagttgttgttggtaaaagagctgaagaaggtattgttgaggtaaaa
     caacgcattaacggtttaagtgaagaagtgcaaattgatgaattagagtattacttacaa
     gaattatttaagaatattaagtaa

     Sequence 3184
45   MKQSKVFIPTMREVPAEAEALSHRLLLKAGLIKQSTSGIYSYLPLATRVLNNISKIIREE
     MESIDAVEILMPALQQAELWEESGRWSAYGPELMRLKDRNGREFALGPTHEEVVTSLVRD
     ELKSYKQLPLTLFQIQSKYRDEKRPRFGLLRGREFLMKDAYSFHSDEASLDATYQDMYQA
     YSRIFKRVGINARPVVADSGAIGGSHTHEFMALSEIGEDTIVYSNESDYAANIEKAEVVY
     HPSHKHSALAELTKVETPNVKTAQEVAEYLKRPLDEIVKTMIFKIDGEFIMFLVRGHHEL
50   NEVKLKSYFGTEHVEMATPDEIVNLVDANPGSLGPIFDKDIKIYADNYLQDLNNFVVGAN
     EDHYHYINVNIGRDFDVTEYGDFRFITQGEMLSDGSGVAQFAEGIEVGQVFKLGTKYSES
     MNATFLDNQGKAQPLIMGCYGIGVSRTLSAIVEQNNDENGIIWPKSVTPFDIHLITINPK
     KDDQRTLGDQLYQKLMDSYDVLYDDRKERAGVKFNDSDLIGLPVRVVVGKRAEEGIVEVK
     QRINGLSEEVQIDELEYYLQELFKNIK*
55
     Sequence 3185
     Contig_0703_pos_2091_3398
     is similar to (with p-value 3.0e-36)
      >sp:sp|Q57991|AK_METJA PROBABLE ASPARTOKINASE (EC 2.7.2.4)

(ASPARTATE KINASE). >gp:gp|U67506|U67506_4 Methanococcus jan
naschii section 48 of 150 of the complete genome. NID: g1591
274.

gtgaatgaagaccctgaacgtaaaatcattatcgtttcagctccaggcaaaaggcataat
gacgacattaaaactactgatttattaattcgtctctatgaaaaagtacttaataaatta
aattatgaaagtaaaaaacaagaaattatccaaagatatgctgatatagtagaagaatta
ggtataggaaatgacattttaataacaattaatgacactttagaggaatacattaaacat
ctttctgacaaacctaaccgtttatatgatgctttattatcttgtggcgaaaattttaat
gctcaattaatagcccagtataataatagtcaaggtattcctactcgttatatttctcct
aaagaagctggattaactgtaactgatttaccacagcaagctcaattttagattccgca
tataatgaaatatacaaattgcgtgattatgatgaaaagctaattattcctggttttttc
ggagtttcaaagcaaaattatatcgttacgtttccacgcggtggttctgacataactggt
gctatcatagcacgtggcgtccgagcctcactttatgagaacttcactgatgtatcagga
atatataaagctaatccgaatatcataaataatcctgaactcatagaggaaataacttat
agagaaatgcgagagctatcttatgcaggatttggagttttttcacgatgaagctctacaa
cctttatacaaagatcgaattcccgtagttatcaaaaatactaatcgtccaaatgataaa
gggacctacattttacatgaccgtgaaatcgattctaaaaatgtcattagtggaattagt
tgtgataaaggctttactgtgattaatattaaaaaatatttaatgaatagattagttgga
tttacacgaaagattcttggcgttttagaagaatttaatatatcatttgaccacatgcct
tctggtattgataacataagtattatcatgcgtacaaatcaaattcaaggtaaagaaagt
caagttcttaatgccatacgcaaacgttgtgaagttgatgaattaagtatcgaccatgat
ttagcagtactaatgattgttggtgaaggtatgaatcaagttgttggtacagctagtaaa
attactcacgcccctttcagaatcaaacattaatttaataatgattaaccaaggtgcttct
gaaatttcaatgatgtttggaattcatgaagcagatgctgaaaaagcagtattatctacg
tacgaattttgttacaacggtgtttgtttaaaaaatttgtgtaaataa

Sequence 3186
VNEDPERKIIIVSAPGKRHNDDIKTTDLLIRLYEKVLNKLNYESKKQEIIQRYADIVEEL
GIGNDILITINDTLEEYIKHLSDKPNRLYDALLSCGENFNAQLIAQYNNSQGIPTRYISP
KEAGLTVTDLPQQAQILDSAYNEIYKLRDYDEKLIIPGFFGVSKQNYIVTFPRGGSDITG
AIIARGVRASLYENFTDVSGIYKANPNIINNPELIEEITYREMRELSYAGFGVFHDEALQ
PLYKDRIPVVIKNTNRPNDKGTYILHDREIDSKNVISGISCDKGFTVINIKKYLMNRLVG
FTRKILGVLEEFNISFDHMPSGIDNISIIMRTNQIQGKESQVLNAIRKRCEVDELSIDHD
LAVLMIVGEGMNQVVGTASKITHALSESNINLIMINQGASEISMMFGIHEADAEKAVLST
YEFCYNGVCLKNLCK*

Sequence 3187
Contig_0707_pos_8234_4395
 >gp:gp|AF067776|AF067776_1 Abiotrophia defectiva extracellu
lar matrix binding protein (emb) gene, partial cds. NID: g32
49002.
gtgaaatctgaagctagacaagcagtacagaataaagcaaatgaacagattaatcatatt
caaaacacgcctgatgcaactaatgaagaaaaacaagaggcaataaatagagtaagtgct
gaattagcaagagttcaagcacaaataaatgcagaacatacaacccaaggtgtcaaaact
atcaaagacgacgcgataacttctttatctcgaattaatgcacaagttgttgagaaagag
tctgcaagaaatgcaatcgaacaaaaggcaacacaacaaacgcaatttattaataataat
gataatgctacagatgaagaaaaagaggtcgccaacaatttagttatcgctacaaaacaa
aaatcattagataatattaactccttatcttcaaataatgatgttgaaaatgctaaagta
gcaggaataaatgaaatagctaacgttttaccagcaaccgctgttaagtcaaaagcaaaa
aaagatattgatcaaaaactcgcgcaacagattaatcaaattcaaacgcatcaaactgct
acaactgaggaaaaagaagcggctattcaattggcaaatcaaaatcaaatgaagcaaga
acagcaattcaaaatgaacatagtaacaatggtgtcgcacaagctaaatctaacggcatt
catgaaattgaattagttatgccagatgcgcacaaaaaatctgatgctaaacaaagtatc
gataataaatataatgagcaaagtaatactatcaacactacaccagatgcaacagatgaa
gaaaagcaaaaagcattagataaattaaaaatagctaaagatgcaggatacaacaaagtt
gatcaagcgcaaacaaaccaacaagtatctgatgcaaaaactgaggctatagatacgata
actaatattcaagcaaatgttgcaaaaaaaaccatccgctcgagtggaattagattcaaag
tttgaggatttaaagcgtcaaatcaatgcaacgcccaatgctacagaagaagaaaaacaa
gatgcaattcaaagattgaatggtaaaagagatgaagttaagaatctaataaatcaagat

797

```
agacgtgacaatgaagttgaacagcacaaaaatattggacttcaagaattagaaacgatt
catgctaatccaactagaaaatctgatgcgctccaagagttacaaactaaatttatttca
caaacagagttaattaataataacaaagatgcaactaatgaagaaaaagatgaagccaaa
cgacttcttgagattagtaaaaataaaactataacaaatatcaatcaagcgcaaactaat
aatcaagttgataatgctaaagataacggcatgaatgagattgctaccataataccagca
acaacaattaaaacagatgcaaaaacggctattgataaaaaagctgagcaacaagttaca
atcatcaatggtaacaacgatgcaacagatgaagaaaaagcagaggctagaaagctggtt
gaaaaagcgaaaattgaagccaaatctaatattacaaatagtgatactgaaagggaagtc
aatggtgctaaaaccaatgggttagaaaaaataaacaatattcaaccatcaactcaaact
aaaacaaatgctaagcaagaaatAaatgacaaagctcaagaacaattaatccaaattaat
aacacgcctgatgcaaccgaagaagaaaagcaagaggcaacaaatagagtcaatgctgga
ttagcacaagcaatacaaaatattaataatgcacatagtactcaagaagtaaatgaatct
aaaacaaatagtattgctacaatcaagagtgtacaacccaatgtgatcaaaaaaccgact
gctataaatagtttgactcaagaagctaataatcaaaagacgttaataggtaatgatggt
aatgctactgatgatgaaaaagaggctgcaaagcaattagtgacccaaaaattaaatgaa
caaattcaaaaaattcatgaaagtacacaagataatcaagttgataacgtaaaagcacaa
gctatcactgcaattaaattgattaatgcaaatgcacataaaagacaagatgccattaat
attttgactaatctagctgaaagtaaaaaatcagatataagagccaatcaagatgcaact
actgaagagaaaaatacggcaatacaatctatagatgatacgttagcacaagcacgtaac
aatattaatggtgcaaatacaaatgcgttagtggatgagaatttagaagatggtaagcaa
aagttacaacgtattgtgttgtcaactcaaactaaaacacaagctaaagcagacattgct
caagcaataggtcaacaaaggtcgacaatagaccagaatcaaaatgctacaacagaagaa
aaacaagaagcccttgagagacttaatcaagaaacaaatggagtcaatgatagaatacaa
gcagctttagcaaatcaaaatgttacagacgaaaaaaataatatattagaaacaataaga
aatgttgaacctattgtaattgtaaaaccaaaggctaatgaaataattagaaaaaaagct
gcggaacaaacgactttaataaatcaaaatcaagatgcgacactagaagaaaaacaaata
gcacttggcaaattagaagaagtaaagaatgaagcgttaaatcaagtatcacaggcacac
tcaaataatgatgtgaaaattgtggaaaataatggaattgctaaaatttctgaggtccat
cctgagactataattaaacgtaatgctaaacaagaaattgaacaagatgcgcaaagtcaa
attgatactatcaatgcaaataataaatcaactaatgaagaaaaatcagccgctatagat
agagttaatgtagctaaaattgatgctattaacaatattactaatgctacaactacacaa
ttagttaatgatgctaaaaatagtggtaacacgagtattagccaaatattaccaagtaca
gcagtcaaaactaatgcattagcagctctagctagcgaagctaaaaataaaaacgctata
atagatcaaacaccaaatgcgacagcagaagaaaaagaagaagcaaataataaagttgat
cgtcttcaagaagaagcagatgctaatatcctaaaagcacacactactgatgaagttaat
aatattaaaaatcaagctgttcaaaatattaacgctgttcaagttgaagttatcaagaaa
caaaacgctaaaaaccaattaaatcaattcattgataatcaaaagaaaattattgaaaat
acgcctgatgcaacactagaagaaaaagctgaagctaatagattgcttcaaaatgtacta
acttccacatcagatgaaattgctaatgtagatcataacaacgaggttgatcaagcttta
gataaagctagaccaaaaatcgaggcaattgtaccacaagttagtaagaaacgagatgct
ttaaatgcaatccaagaagcatttaattcacaaactcaagaaatacaagagaagcaagaa
gctacgaatgaagaaaaaactgaagcattaaataaaaataaaccaattacttaatcaggct
aaagtaaatattgatcaagcacagtcaaataaagatgtagatagtgcgaaaacacgtagt
attcaagatatagagcaaattcaaccacatccacaaacaaaagcaaccgggcgtcacaga
ttaaatgaaaaagctaaccaacaacaaagtactattgcaactcatcctaattcaacaatt
gaagaaagacaggaagcaagtgcaaaactacaagaagttcttaaaaaaagccatagctaa
```

Sequence 3188
VKSEARQAVQNKANEQINHIQNTPDATNEEKQEAINRVSAELARVQAQINAEHTTQGVKT
IKDDAITSLSRINAQVVEKESARNAIEQKATQQTQFINNNDNATDEEKEVANNLVIATKQ
KSLDNINSLSSNNDVENAKVAGINEIANVLPATAVKSKAKKDIDQKLAQQINQIQTHQTA
TTEEKEAAIQLANQKSNEARTAIQNEHSNNGVAQAKSNGIHEIELVMPDAHKKSDAKQSI
DNKYNEQSNTINTTPDATDEEKQKALDKLKIAKDAGYNKVDQAQTNQQVSDAKTEAIDTI
TNIQANVAKKPSARVELDSKFEDLKRQINATPNATEEEKQDAIQRLNGKRDEVKNLINQD
RRDNEVEQHKNIGLQELETIHANPTRKSDALQELQTKFISQTELINNNKDATNEEKDEAK
RLLEISKNKTITNINQAQTNNQVDNAKDNGMNEIATIIPATTIKTDAKTAIDKKAEQQVT
IINGNNDATDEEKAEARKLVEKAKIEAKSNITNSDTEREVNGAKTNGLEKINNIQPSTQT
KTNAKQEINDKAQEQLIQINNTPDATEEEKQEATNRVNAGLAQAIQNINNAHSTQEVNES
KTNSIATIKSVQPNVIKKPTAINSLTQEANNQKTLIGNDGNATDDEKEAAKQLVTQKLNE
```

QIQKIHESTQDNQVDNVKAQAITAIKLINANAHKRQDAINILTNLAESKKSDIRANQDAT
TEEKNTAIQSIDDTLAQARNNINGANTNALVDENLEDGKQKLQRIVLSTQTKTQAKADIA
QAIGQQRSTIDQNQNATTEEKQEALERLNQETNGVNDRIQAALANQNVTDEKNNILETIR
NVEPIVIVKPKANEIIRKKAAEQTTLINQNQDATLEEKQIALGKLEEVKNEALNQVSQAH
SNNDVKIVENNGIAKISEVHPETIIKRNAKQEIEQDAQSQIDTINANNKSTNEEKSAAID
RVNVAKIDAINNITNATTTQLVNDAKNSGNTSISQILPSTAVKTNALAALASEAKNKNAI
IDQTPNATAEEKEEANNKVDRLQEEADANILKAHTTDEVNNIKNQAVQNINAVQVEVIKK
QNAKNQLNQFIDNQKKIIENTPDATLEEKAEANRLLQNVLTSTSDEIANVDHNNEVDQAL
DKARPKIEAIVPQVSKKRDALNAIQEAFNSQTQEIQEKQEATNEEKTEALNKINQLLNQA
KVNIDQAQSNKDVDSAKTRSIQDIEQIQPHPQTKATGRHRLNEKANQQQSTIATHPNSTI
EERQEASAKLQEVLKKSHS*

Sequence 3189
Contig_0711_pos_3905_0
is similar to (with p-value 2.0e-53)
 >sp:sp|O69282|MQO_CORGL MALATE:QUINONE OXIDOREDUCTASE (EC 1
.1.99.16) (MALATE DEHYDROGENASE (ACCEPTOR)) (MQO).
gtgagcaaaaaaatggctaataaagagtcaaaaaatgttgttattattggcgctggtgtc
ttaagtacgacatttggttctatgattaaagaattagaacctgattggaacatcaaactc
tatgaacgcttagatcgtccaggtattgaaagttctaacgaaagaaacaatgccggtaca
ggacatgcggcgttatgtgaattgaactatacagtacaacaacctgatggttcaattgat
atagaaaaagccaaagaaatcaacgaacaattcgagatttcaaaacaattctggggtcac
ttagtaaaaagtggtaacatcagtaaccctagagatttcattaatccacttcctcacatt
agtttcgtaagaggtaaaaataacgttaaattcttaaaaaaccgttacgaagcaatgcgt
aacttccctatgttcgataacatcgaatatacagaagatatcgaagaaatgagaaaatgg
atgccattaatgatgacaggtcgtactggtaacgaaatcatggcggctagtaaaatcgac
gaaggtacagatgttaactacggtgaattaactcgtaaaatggcaaaaagtattgaaaaa
catccaaatgctgatgttcaatacaaccacgaagtaattaatttcaatcgtcgtaaagac
ggtatttgggaagttaaagtta

Sequence 3190
VSKKMANKESKNVVIIGAGVLSTTFGSMIKELEPDWNIKLYERLDRPGIESSNERNNAGT
GHAALCELNYTVQQPDGSIDIEKAKEINEQFEISKQFWGHLVKSGNISNPRDFINPLPHI
SFVRGKNNVKFLKNRYEAMRNFPMFDNIEYTEDIEEMRKWMPLMMTGRTGNEIMAASKID
EGTDVNYGELTRKMAKSIEKHPNADVQYNHEVINFNRRKDGIWEVKVX

Sequence 3191
Contig_0712_pos_4917_4060
is similar to (with p-value 3.0e-91)
 >gp:gp|D13095|BACPK_1 B. stearothermophilus phosphofructoki
nase and pyruvate kinase genes. NID: g285620.
atgtttaaagatttttttaatcgaagcaagaaaaagaaatatttaacagttcaagattct
aaacaaaatgatgtacctgctggtataatgacaaaatgtcctaattgcaaaaaaataatg
tatacaaaagaattgaatgaaaatttaaatgtatgctttaattgtgatcatcatatagct
ttaaccgcatataaaagaatagaagcaatttcagacgatggatcatttatagaatttgat
agaggtatgacatctgctaacccattagactttcctgggtatgaagaaaaaattgaaaaa
gatcagcaaaagactggacttaatgaagcgttagtgtctggtactgcgaaattagatgga
atacaatatggtgttgcagttatggatgctcgttttagaatgggaagcatgggctctgta
gttggtgaaaaaatatgcagaattattgattattgtacagaacatcgtttgccatttatt
ctgttttctgcgagtggtggagctagaatgcaagagggaattatttctttaatgcaaatg
gggaaaacaagtgtttctttaaaaagacattctgatgcaggactattatatattcttac
ataactaatcccactactggagggggtttctgcaagtttgcttcggttggagatattaat
ttaagtgaacctaaagcactaatcggatttgctggtagacgtgttatagaacaaacaatt
aatgaaaagttgcctgatgatttccaaactgctgagtttttattagagcatggtcaactt
gataaagtcattcatcgaaaagatatgcgtgagactttatcaaatattttaaaaatccat
caagaggtgagtaactaa

Sequence 3192
MFKDFFNRSKKKKYLTVQDSKQNDVPAGIMTKCPNCKKIMYTKELNENLNVCFNCDHHIA

LTAYKRIEAISDDGSFIEFDRGMTSANPLDFPGYEEKIEKDQQKTGLNEALVSGTAKLDG
IQYGVAVMDARFRMGSMGSVVGEKICRIIDYCTEHRLPFILFSASGGARMQEGIISLMQM
GKTSVSLKRHSDAGLLYISYITNPTTGGVSASFASVGDINLSEPKALIGFAGRRVIEQTI
NEKLPDDFQTAEFLLEHGQLDKVIHRKDMRETLSNILKIHQEVSN*

Sequence 3193
Contig_0712_pos_3958_3116
is similar to (with p-value 2.0e-74)
>sp:sp|Q54776|ACCD_SYNP7 ACETYL-COENZYME A CARBOXYLASE CARB
OXYL TRANSFERASE SUBUNIT BETA (EC 6.4.1.2). >gp:gp|U59237|SP
U59237_6 Synechococcus PCC7942 ORF102, ORF120, ORF113, ORF12
8, CTP synthetase (pyrG), carboxyltransferase beta subunit (
accD), ORF145 and ORF123 genes, complete cds. NID: g1399849.

atgctggaagcatcattaaaaagagaaactacaaaagtgtacactaatctaaaaccttgg
gatcgtgttcaaatcgctcgtttaccagaaagaccaaccacattagattatattccctat
atttttgattcatttattgagttacatggcgatagaagtttttagggatgatccagcaatg
attggtggaattggttacttagatggtaagtctgtaacagttataggccaacaacgtggt
aaagacacgaaagataaatattatcgtaattttggtatggctcacccagaagggtataga
aaagctttgcgtttaatgaaacaagcagagaaatttaatcgtccaatatttacttttata
gatactaaaggtgcttatccgggtaaagcagctgaagaaagaggtcaaagtgaatcaatt
gcaaaaaattgatggaaatggcttcattaacggtaccagttattgctgttgttattggt
gaaggcggaagtggcggcgctttaggaattggaatctcaaatcgtgttctgatgcttgaa
aatagtacttattcagttatttcacctgaaggagcagctgacttttatggaaagatagt
aacttagctcaaattgccgctgaaactatgaaaatcactgcgcatgatttactagattta
ggtattatagatgaagtgattaatgagccacttggtggtgcgcaaaaagatgaagaagca
caagctttatcaattaagaaaatgttccttaaacatttaaatgaattaaagcaactcaca
cctgaagaattagcaaatgatcgttttgaaaaatttagaaaaattggttcagttgtggag
tga

Sequence 3194
MLEASLKRETTKVYTNLKPWDRVQIARLPERPTTLDYIPYIFDSFIELHGDRSFRDDPAM
IGGIGYLDGKSVTVIGQQRGKDTKDNIYRNFGMAHPEGYRKALRLMKQAEKFNRPIFTFI
DTKGAYPGKAAEERGQSESIAKNLMEMASLTVPVIAVVIGEGGSGGALGIGISNRVLMLE
NSTYSVISPEGAAALLWKDSNLAQIAAETMKITAHDLLDLGIIDEVINEPLGGAQKDEEA
QALSIKKMFLKHLNELKQLTPEELANDRFEKFRKIGSVVE*

Sequence 3195
Contig_0714_pos_982_1668
is similar to (with p-value 9.0e-54)
>sp:sp|P54374|AROE_BACSU SHIKIMATE 5-DEHYDROGENASE (EC 1.1.
1.25). >gp:gp|D84432|BACJH642_93 Bacillus subtilis DNA, 283
Kb region containing skin element. NID: g2627063. >gp:gp|Z99
117|BSUB0014_46 Bacillus subtilis complete genome (section 1
4 of 21): from 2599451 to 2812870. NID: g2634966.
atgataggaattattggagcaatggaagaagaagtgacgattttaaagcgtaaattgaat
gatatgaatgaaataaatattgcgcatgttaaattttatgttggcaagctaaaccacaaa
gaggtggttttaacacaaagtggtataggtaaagttaatgcttctatctcaacgactttg
ttaatagaaaaatttaatccagaagtcgtcattaatactggatcagcaggtgcactagat
caaacactatctattggagatatattagtgagtaatcatgtattatatcatgatgctaat
gctacagcgtttggttatgaatatggacaaatacctcaaatgcctaaaacttatactact
gatcctactttgttgaaaaaaacaatgcatgtattagaacaacaacaactgaatggtaaa
gtaggtatgattgttagtggtgatagtttataggtagctcagaacagcgacaaaaaatt
aagcaacaattccagaagctatggctgtcgaaatggaggcaactgcaattgcgcaaaca
tgttatcaatttaaagtaccattatcgtaactagagctgtttctgatttagcaaacggt
aaagccgatatttcttttgaagaatttttagataaagcagctttatcatctagtgagaca
gtttcattattagtagaatcattataa

Sequence 3196

MIGIIGAMEEEVTILKRKLNDMNEINIAHVKFYVGKLNHKEVVLTQSGIGKVNASISTTL
LIEKFNPEVVINTGSAGALDQTLSIGDILVSNHVLYHDANATAFGYEYGQIPQMPKTYTT
DPTLLKKTMHVLEQQQLNGKVGMIVSGDSFIGSSEQRQKIKQQFPEAMAVEMEATAIAQT
CYQFKVPFIVTRAVSDLANGKADISFEEFLDKAALSSSETVSLLVESL*

Sequence 3197
Contig_0714_pos_3323_4141
is similar to (with p-value 1.0e-57)
  >sp:sp|P24247|PFS_ECOLI PFS PROTEIN (P46). >pir:pir|S45227|
  S45227 purine nucleoside phosphorylase homolog - Escherichia
  coli >gp:gp|D26562|ECO82K_47 Escherichia coli genome, 2.4-4
  .1 min region (110,917-193,643 bp from 0 min). NID: g473770.
  >gp:gp|U70214|ECU70214_10 Escherichia coli chromosome minut
  es 4-6. NID: g1552727. >gp:gp|AE000125|AE000125_6 Escherichi
  a coli K-12 MG1655 section 15 of 400 of the complete genome.
  NID: g1786348. >gp:gp|U24438|ECU24438_1 Escherichia coli MT
  A/SAH nucleosidase gene, complete cds. NID: g2981266.
  gtgataaaagtgaaatttgcagtaattggaaacccatttctcattcattatcgccattg
  atgcatcatgctaattttcaatctttaaatttggaaaacacgtatgaagcgataaatgta
  ccagttaatcaatttcaagacattaaaaaaataatttcagaaaagagtattgatggattc
  aatgttactattccacataaagaacgtattattccgtacctagatgatattaatgaacaa
  gcgaaatctgttggggcggtaaatacagtttttagttaaagatggtaagtggattggttat
  aatactgatggaattggttatgttaatggtttaaaacaaatatatgaaggtatagaagac
  gcttatatattaattttaggtgcaggtggagcaagtaaaggtatagcaaatgaattatat
  aaaatcgttcgtccgactttaacagttgcaaatagaacgatgtctcgtttttaataattgg
  tcgttaaatattaacaaaataaatttaagccatgcagaaagccatttagatgaatttgat
  attataataaacactacacctgctggtatgaacggcaatacagattctgtaatttctttta
  aatcgtttagcttcacatactttagtaagtgatattgtttataatccatataaaacacca
  atactaatagaagctgaacaaagaggtaatccaatctataatggtcttgatatgttcgtt
  catcaaggtgctgaaagtttttaaaatttggactaatctagaaccagatataaaagcaatg
  aaaaacatagtaattcaaaaattgaaaggagaattatga

Sequence 3198
VIKVKFAVIGNPISHSLSPLMHHANFQSLNLENTYEAINVPVNQFQDIKKIISEKSIDGF
NVTIPHKERIIPYLDDINEQAKSVGAVNTVLVKDGKWIGYNTDGIGYVNGLKQIYEGIED
AYILILGAGGASKGIANELYKIVRPTLTVANRTMSRFNNWSLNINKINLSHAESHLDEFD
IIINTTPAGMNGNTDSVISLNRLASHTLVSDIVYNPYKTPILIEAEQRGNPIYNGLDMFV
HQGAESFKIWTNLEPDIKAMKNIVIQKLKGEL*

Sequence 3199
Contig_0717_pos_7272_6145
is similar to (with p-value 1.0e-99)
  >sp:sp|P54524|YQIG_BACSU PROBABLE NADH-DEPENDENT FLAVIN OXI
  DOREDUCTASE YQIG (EC 1.-.-.-). >gp:gp|D84432|BACJH642_230 Ba
  cillus subtilis DNA, 283 Kb region containing skin element.
  NID: g2627063. >gp:gp|Z99116|BSUB0013_132 Bacillus subtilis
  complete genome (section 13 of 21): from 2395261 to 2613730.
  NID: g2634723.
  atgaataataaatatgaacctttatttaaatctttaacactacctaatggtgttgaagta
  agaaatcgttttgtttttagctcctcttacacatacttcatcaaatgatgatggaacaatt
  tcagatatagaattaccttacattgagaaacgttctaaagatgttgggattgcaattaat
  gcagcaagtaatgttaatgatgtaggcaaagcatttcctggtcaaccttctgttgcacat
  gactcagatatcgaagggttaaaagaacttgctcaagttatgaagaaaaatggtgcgaaa
  gcaatagtccaaattcatcatggaggtgctcaagctttaccagagttaacacctgatgga
  gatgttgttgcaccaagtgccatttctcttaaaagtttttggtcagcaaaaagaacatgat
  gctcgtgagatgactgctgaagaaattgaacaaactattagagactttggtgaagcaact
  agaagagcaattgaagcaggttttgatggcgttgaaattcatggcgcaaaccattatctt
  attcaccaatttgtttctccttactataatagaagaaatgatgttgggctgataactat
  aaattccctgttgctgttatagatgaagttgttaaagctaaaaaagctcatgcatatgat

```
gattttattattggttacagattgtcacctgaagaagcggaatcaccaggtatttcaatg
gagataactgaagaattaattcaccaaatcgcaaataaaccacttgattatattcatgtg
tcattaatggatgttaactcagttacgcgagaaggtaaatataaaggtgaaaatcgcttg
gaacttattcatcaatggataaatggacgtatgccgcttattggtataggttctgtcttt
```
5  
```
acagctgaagatgcactaaatgctgttgaaaacattggagttgaatttgttgcgttaggt
tgtgaaattctacttgattatgattttgttgctaaaattaaagaaggtcgagaagacgaa
attataaatgcttttgatcctaatcgtgaagaccaacattatctaacaccaaatctttgg
gaacagtttaatcaaggattctatccattacctcgaaaagacaaataa
```

10  Sequence 3200  
```
MNNKYEPLFKSLTLPNGVEVRNRFVLAPLTHTSSNDDGTISDIELPYIEKRSKDVGIAIN
AASNVNDVGKAFPGQPSVAHDSDIEGLKELAQVMKKNGAKAIVQIHHGGAQALPELTPDG
DVVAPSAISLKSFGQQKEHDAREMTAEEIEQTIRDFGEATRRAIEAGFDGVEIHGANHYL
IHQFVSPYYNRRNDVWADNYKFPVAVIDEVVKAKKAHAYDDFIIGYRLSPEEAESPGISM
```
15  
```
EITEELIHQIANKPLDYIHVSLMDVNSVTREGKYKGENRLELIHQWINGRMPLIGIGSVF
TAEDALNAVENIGVEFVALGCEILLDYDFVAKIKEGREDEIINAFDPNREDQHYLTPNLW
EQFNQGFYPLPRKDK*
```

Sequence 3201  
20  Contig_0718_pos_2272_1247  
```
is similar to (with p-value 3.0e-29)
 >sp:sp|P31547|YAEE_ECOLI HYPOTHETICAL ABC TRANSPORTER PERME
ASE PROTEIN YAEE. >gp:gp|D83536|ECOTSF_23 Escherichia coli g
enome, 4.0 - 6.0 min region. NID: g1208942. >gp:gp|U70214|EC
```
25  ```
U70214_47 Escherichia coli chromosome minutes 4-6. NID: g155
2727. >gp:gp|AE000129|AE000129_2 Escherichia coli K-12 MG165
5 section 19 of 400 of the complete genome. NID: g1786395.
gtgattgaattcaaaaatgttaacaaagtttttcgcaaaaaaagagaaactattcaagct
ttgaaaaatgtatcatttaagattgaccaacatgatattttggtgttattggttatagt
```
30 ```
ggtgctggtaaaagtacattagttcggttagtcaatcaacttgagacagtatcagatggt
caagttattgttgatggtcatgagattgatacatataaagaaaaagatttacgtgatatt
aaaaaagatatcggtatgatctttcaacatttcaatttgcttaattctaaatcagtctat
aaaaatgttgcaatgccacttattttaagtaagacaaataagaaagaaattaaggaaaaa
gttgacgaaatgttagaatttgtggggcttgctgataaaaaagatcaatttccagatgaa
```
35  ```
ttatcaggtggacaaaaacaacgtgttgccatcgcaagagcattagtaacgcatcctaaa
atattattatgtgatgaagcgacaagtgctctggatccagctactacaagctcaatttta
aatttattaagtaatgtgaatcgaacatttggtgtgacgattatgatgattacacatgaa
atgagcgtaattcaaaaaatttgtcatcgtgtagctgtcatggaaatggcgaagtgata
gaaatggggacagttaaagatgtctttagtcatccacaaacgaacactgcaaaaaatttc
```
40 ```
gtttcgacggtgattaacactgagccttcaaaagagttacgggcctcttttaactcgaga
aaagattcaaatttcacagattataaactgtttttagactctgaacaaattcaattgcca
atattgaacgagcttatcaacgagcatcatcttaacgttaacgtattattttcttctatg
tcagaaattcaagatgaaacggtttgttatttgtggttgagatttgagcatgatgagtca
tttaatgatttttaaacttactgattacctttcaaaacgacatattcggtatgaggaggtt
```
45  ```
atataa
```

Sequence 3202
```
VIEFKNVNKVFRKKRETIQALKNVSFKIDQHDIFGVIGYSGAGKSTLVRLVNQLETVSDG
QVIVDGHEIDTYKEKDLRDIKKDIGMIFQHFNLLNSKSVYKNVAMPLILSKTNKKEIKEK
```
50 ```
VDEMLEFVGLADKKDQFPDELSGGQKQRVAIARALVTHPKILLCDEATSALDPATTSSIL
NLLSNVNRTFGVTIMMITHEMSVIQKICHRVAVMENGEVIEMGTVKDVFSHPQTNTAKNF
VSTVINTEPSKELRASFNSRKDSNFTDYKLFLDSEQIQLPILNELINEHHLNVNVLFSSM
SEIQDETVCYLWLRFEHDESFNDFKLTDYLSKRHIRYEEVI*
```

55  Sequence 3203  
Contig_0718_pos_1244_585  
```
is similar to (with p-value 6.0e-81)
 >sp:sp|P30750|ABC_ECOLI ATP-BINDING PROTEIN ABC. >gp:gp|U70
214|ECU70214_48 Escherichia coli chromosome minutes 4-6. NID
```

: g1552727. >gp:gp|AE000129|AE000129_3 Escherichia coli K-12
MG1655 section 19 of 400 of the complete genome. NID: g1786
395.

atgtttggttcaagtttagattcatctcaattattacaagctctatacgaaacattgtat
atggtgactgtatcacttgtaatcggtgctttaataggtatacctcttggcatcttgtta
gtggtaactagaaaaaacggtatatggtcgaatacaatattgcatcaagtattaaatcct
atcattaatattttaagatcaattccgttcattattttattaatagccatagtgcctttt
actaaattgctagttggcacatctatcggcacaacagcagccatagtaccactcacggtt
tatgtagcaccttatatcgcacgcttagtagaaaactcattactggaagtcgatgacggg
attattgaggcagctaaagcaatgggtgcatcacctcttcaaattatacgttattttttta
ttgccagaagcacttggttcattaattctagctataactacagctattattggtctcata
ggtagtacagcaatggctggtgctgttggtggtggcggtataggtgatttggctttagtg
tatggttatcaacgattcgatacaattgtcattgtgattacagtcattgtacttattatt
attgttcaaattatacaaacgttaggtaactttatcgctagggttatccgtagaaattaa

Sequence 3204
MFGSSLDSSQLLQALYETLYMVTVSLVIGALIGIPLGILLVVTRKNGIWSNTILHQVLNP
IINILRSIPFIILLIAIVPFTKLLVGTSIGTTAAIVPLTVYVAPYIARLVENSLLEVDDG
IIEAAKAMGASPLQIIRYFLLPEALGSLILAITTAIIGLIGSTAMAGAVGGGGIGDLALV
YGYQRFDTIVIVITVIVLIIIVQIIQTLGNFIARVIRRN*

Sequence 3205
Contig_0719_pos_4886_3567
>sp:sp|P52673|CYSI_THIRO SULFITE REDUCTASE (NADPH) HEMOPROT
EIN BETA-COMPONENT (EC 1.8.1.2) (SIR-HP). >gp:gp|Z23169|TRCY
SCOMA_3 T.roseopersicina cysJ, cysI, cysH genes, complete CD
S, and cysB gene 5' end. NID: g1518424.
atgaaaaatattaatcatgcagtacttgattctattgctgcatgtggagatgttaatcgt
aatacgatgtgcaatcctaatccttatcaatctcaagtacataaggagattaatgattat
gcaacgcgtataagtaatcacttacttccaagaacaaatgcatatcatgaaatttggctt
gatggtgaaaaggttttagattcgagtgaggaaaaggaacctatttatgggaatacgtat
ttaccacgtaaattcaaaataggtattgcagtaccaccatctaatgatattgacgtctat
tctcaagatattggtttaatcgctatcgttgaacaagatgagttaattggatttaatgtg
actatcggtggcggtatgggtatgactcatggtaatactgaaacatatcctcaacttgga
cgtctcataggtttttatacctaaggaaaaggttgtagatgtatgtgagaaaatacttaca
atacaacgtgattatggtaatcgtgaaatcgaaaaaatgcacgtttttaaatatacagtg
gaccgtctaggagaaacttgggtgactgaagaattaaaccgacgattaggttgggaaatt
aaagcgccacgtgatttcgaatttgaacataatggtgatcgattaggttggattgaaggt
attaataattggaatttcactttatttatacaaaatgggcgtgtgaaagatactgaagac
tatttgttaaaaacaaccttaagagaaatcgcagaaatccatactggagatttcagatta
tcacctaatcagaacttagttattgcaaatgtttctcctgagaaaaaggaagaaatacaa
gctattattgataaatataaattaacagatggcaaaaattatacaggacttagaagaaat
tctatggcttgtgttgctttcccaacgtgtggtttagctatggcagaatctgaaagatat
cttccttcactaattacaaaaattgaagattattagatgagtctggtttaaaagaggaa
gaaataacgattcgtatgacaggttgtcccaatggatgtgcgagaccagcgctagcagaa
atagcctttatcggtaaagcacctggtaaatataatatgtacttaggtggtagttttaaa
ggcgaacgtctaaataaaatatataaagagaatatcgacgaaatgagatattagaaagt
ctacgtccattgttgttgcgttatagtaaagagcgtcttgacggagaacactttggggac
tttgtaattcgtgacggtgtgatagccaaagttcatgatggtcgcgatttcatagttaa

Sequence 3206
MKNINHAVLDSIAACGDVNRNTMCNPNPYQSQVHKEINDYATRISNHLLPRTNAYHEIWL
DGEKVLDSSEEKEPIYGNTYLPRKFKIGIAVPPSNDIDVYSQDIGLIAIVEQDELIGFNV
TIGGGMGMTHGNTETYPQLGRLIGFIPKEKVVDVCEKILTIQRDYGNRENRKNARFKYTV
DRLGETWVTEELNRRLGWEIKAPRDFEFEHNGDRLGWIEGINNWNFTLFIQNGRVKDTED
YLLKTTLREIAEIHTGDFRLSPNQNLVIANVSPEKKEEIQAIIDKYKLTDGKNYTGLRRN
SMACVAFPTCGLAMAESERYLPSLITKIEDLLDESGLKEEEITIRMTGCPNGCARPALAE
IAFIGKAPGKYNMYLGGSFKGERLNKIYKENIDENEILESLRPLLLRYSKERLDGEHFGD
FVIRDGVIAKVHDGRDPHS*

Sequence 3207
Contig_0728_pos_9983_10327
>gp:gp|AF043386|AF043386_2 Clostridium acetobutylicum glyce
raldehyde-3-phosphate dehydrogenase (gap), phosphoglycerate
kinase (pgk), and triosephosphate isomerase (tpi) genes, com
plete cds; and 2,3-bpg-independent phosphoglycerate mutase (
pgm-i) gene, partial cds. NID: g2829136.
atgcagattgaagaaaccttccgagacttgaaaagtcctgcctacggactaggcctacgc
catagccgaacgagcagctcagagcgtttttgatatcatgctgctaatcgccctgatgctt
caactaacatgttggcttgcgggcgttcatgctcagaaacaaggttgggacaagcacttc
caggctaacacagtcagaaatcgaaacgtactctcaacagttcgcttaggcatggaagtt
ttgcggcattctggctacacaataacaagggaagactcactcgtggctgcaaccctgctt
actcaaaatctattcacacatggttacgttttggggaaattatga

Sequence 3208
MQIEETFRDLKSPAYGLGLRHSRTSSSERFDIMLLIALMLQLTCWLAGVHAQKQGWDKHF
QANTVRNRNVLSTVRLGMEVLRHSGYTITREDSLVAATLLTQNLFTHGYVLGKL*

Sequence 3209
Contig_0728_pos_8652_7642
is similar to (with p-value 4.0e-59)
>pir:pir|I67760|I67760 transposase (insertion sequence IS10
) - Escherichia coli >gp:gp|S67119|S67119_2 BST=somatotropin
...BST/beta-Gal fusion protein [Escherichia coli, LBB84, pla
smid pXT107, IS10L/R-1, PlasmidTransposonInsertionMutant, 3
genes, 1679 nt]. NID: g455674.
atggcaattaaagtagcaattaatggttttggtagaattggtcgtttagcattcagaaga
attcaagatgtagaaggtcttgaagtagttgcagttaacgacttaacagatgacgatatg
ttagctcatttattaaaatacgatactatgcaaggtcgtttcactggagaagttgaagtt
atcgaaggtggattccgtgtgaacggtaaagaaattaaatcattcgatgaaccagatgct
ggtaaattaccatggggcgatttagatatcgacgtagtattagaatgtactggtttctat
actgataaagaaaaagcacaagctcacatcgatgcaggtgctaaaaaagtattaatctca
gctccagctaaaggtgatgtaaaaacaatcgtattcaacactaaccatgatacattagat
ggttcagaaacagttgtttcaggtgcttcttgtactactaactcattagcaccagttgca
aaagttttaagtgacgaattcggtttagttgaaggtttcatgactacaattcacgcttac
actggtgaccaaaatacacaagacgcacctcacagaaaaggtgacaaacgtcgtgcacgt
gcagcagctgaaaatattatccctaactcaacaggtgctgctaaagctatcggtaaagtt
attccagaaatcgatggtaaattagacggtggagcacaacgtgttccagttgctactggt
tctttaactgaattaactgtagtattagacaaacaagatgtaactgttgaccaagttaac
agtgctatgaaacaagcttctgacgaatcattcggttacactgaagacgaaatcgtatct
tctgatattgttggtatgacttacggttcattatttgatgcgactcaaactcgtgttatg
actgttggagatcgtcaattagttaaagttgcagcttggtacgacaatgaaatgtcttac
actgctcaattagtacgtacattagctcacttagctgaactttctaaataa

Sequence 3210
MAIKVAINGFGRIGRLAFRRIQDVEGLEVVAVNDLTDDDMLAHLLKYDTMQGRFTGEVEV
IEGGFRVNGKEIKSFDEPDAGKLPWGDLDIDVVLECTGFYTDKEKAQAHIDAGAKKVLIS
APAKGDVKTIVFNTNHDTLDGSETVVSGASCTTNSLAPVAKVLSDEFGLVEGFMTTIHAY
TGDQNTQDAPHRKGDKRRARAAAENIIPNSTGAAKAIGKVIPEIDGKLDGGAQRVPVATG
SLTELTVVLDKQDVTVDQVNSAMKQASDESFGYTEDEIVSSDIVGMTYGSLFDATQTRVM
TVGDRQLVKVAAWYDNEMSYTAQLVRTLAHLAELSK*

Sequence 3211
Contig_0732_pos_1186_2424
>gp:gp|AF043609|AF043609_1 Arthrobacter viscosus aluminum r
esistance protein (Alu-2) gene, complete cds. NID: g2827438.

atgcaagattttagcaatttagttgaagaagttgaaaacacacttattccttactttaga

```
     aaaattgaaaagcgtgcattatttaatcaggaaaaggtcttaaatgcttttcaccatgtt
     aaagctagcgaaagtgatttacaggggtctacgggttatggatatgatgattttgggaga
     gaccatttagaacaaatttatgcgcacacatttaaagcagatgacgcacttgtaagacct
     caaattatttcaggtactcatgctattactttagctttacaaagtacgttaaaaaacaat
   5 gatgaactactttatattacaggtagtccatatgatacacttctagaagtcattggtata
     aatggcaatggtgttgaaagtcttaaagaatatggtgttcgctataatgaagtcgaatta
     cgtgacggtcgaattgatattcctaaagtcatcactgcaattaatgacaatacaaaagtt
     gtagcaattcaacgatcaaaaggatatgatcaacgtccatcaattacaattaatgaaatt
     gaacaagcaataacatctattaaagaggtttatcccaatatcattattttgttgataat
  10 tgttatggagaatttgtagaagataaagaaccgattgaagtaggtgctgatttaatcgcc
     ggatcattaattaaaaatccaggtggaggtttagctaaaattggaggatatattgctggt
     agacaagacttaattgaacgctgtggttatcgtttaacagcaccaggcattggtaaggaa
     gcaggagcctcacttaattctttacaagaaatgtatcaaggattcttttctagcgccacat
     gtggttagccaaagtttaaaaggtgcactgtttactagtttgttgttattagaaaaaataaac
  15 atgaagacctcccctaaatataatgtttatcgtacagacttaattcaaacggttcaattt
     gagaccaaagagcaaatgatttcattttgccaaagtatacaacacgcttcaccaattaac
     gcacattttagtccagaacctagctatatgcctggatacgaagatgatgtcatcatggct
     gcaggtacatttattcagggctcgtctattgaattatccgcagacggacctatacgtccg
     ccttatgaagcatatgttcaaggtggtttaacttatgaacatgtcaaattagctgttaca
  20 cgtgcggtgcaacatatgcaagaaaacaatttactataa
```

Sequence 3212
MQDFSNLVEEVENTLIPYFRKIEKRALFNQEKVLNAFHHVKASESDLQGSTGYGYDDFGR
DHLEQIYAHTFKADDALVRPQIISGTHAITLALQSTLKNNDELLYITGSPYDTLLEVIGI
NGNGVESLKEYGVRYNEVELRDGRIDIPKVITAINDNTKVVAIQRSKGYDQRPSITINEI
EQAITSIKEVYPNIIIFVDNCYGEFVEDKEPIEVGADLIAGSLIKNPGGGLAKIGGYIAG
RQDLIERCGYRLTAPGIGKEAGASLNSLQEMYQGFFLAPHVVSQSLKGALFTSLLLEKIN
MKTSPKYNVYRTDLIQTVQFETKEQMISFCQSIQHASPINAHFSPEPSYMPGYEDDVIMA
AGTFIQGSSIELSADGPIRPPYEAYVQGGLTYEHVKLAVTRAVQHMQENNLL*

Sequence 3213
Contig_0732_pos_2905_3270
is similar to (with p-value 5.0e-55)
 >gp:gp|X76490|SAGLNAR_2 S.aureus (bb270) glnA and glnR gene
s. NID: g1134885.
```
     atgtctaatgattcaatcagacgaaacatggccgttttctctatgagtgtggttagtaaa
     ttgacagatttatcaccaagacaaattcgttactatgaaacacatgaacttgtgatgcct
     gaaagaacagatggaaataagagattattttctatgaacgatttagagaggttgttagaa
     ataaagtctcttatcgaaaagggatttaatattagaggtattaaacaaattatattcgat
  40 gagcaagggcatttaactactgatgaacaagagacaagaaagagaatgattgttgacgca
     acgcagaaaccacgtagtgaaacattaccaataaatcgcggcgatttatctcgatttatt
     aaatga
```

Sequence 3214
MSNDSIRRNMAVFSMSVVSKLTDLSPRQIRYYETHELVMPERTDGNKRLFSMNDLERLLE
IKSLIEKGFNIRGIKQIIFDEQGHLTTDEQETRKRMIVDATQKPRSETLPINRGDLSRFI
K*

Sequence 3215
Contig_0733_pos_1000_1344
is similar to (with p-value 2.0e-46)
 >sp:sp|P94453|ALF_BACST FRUCTOSE-BISPHOSPHATE ALDOLASE (EC
4.1.2.13) (FRAGMENT).
```
     atgaaagaaatgttaatcgatgcgaaagaaaacggttatgcggttggtcaatacaatctt
     aataacctcgaatttacacaagctattttagaagcgtctcaagaagagaatgcgccagtt
     attttaggtgtttctgaaggggcagctcgttatatgagtggtttttatacagttgtgaaa
     atggtagaaggtttaatgcatgacttaaacatcacaatcccagtagcaattcatttagac
     cacggttcaagctttgaaaaatgtaaagaagcaattgatgctggattcacatctgtaatg
     attgatgcatctcatagtccttttgaagaaaatgttgaaatatag
```
```

805

Sequence 3216
MKEMLIDAKENGYAVGQYNLNNLEFTQAILEASQEENAPVILGVSEGAARYMSGFYTVVK
MVEGLMHDLNITIPVAIHLDHGSSFEKCKEAIDAGFTSVMIDASHSPFEENVEI*

Sequence 3217
Contig_0733_pos_12244_11498
is similar to (with p-value 4.0e-40)
 >sp:sp|P55476|NODI_RHISN NODULATION ATP-BINDING PROTEIN I.
>gp:gp|AE000076|AE000076_5 Rhizobium sp. NGR234 plasmid pNGR
234a, section 13 of 46 of the complete plasmid sequence. NID
: g2182419.
gtgagtggtcactcattttatcggtgtgaggaggaacaatttatgatagaggtaaaaat
gtaagtaaatcctttggtaaacaacaagtggtagatgatatatctatatcatttaatgct
ggtgaggttgtgggtctaatcggtccttcaggtactggaaaaacgacattaatacagtgc
attttaggcatggagaaaattgatggtgggcaagtcactattcaagaacatacaatgccg
aatagaaagatattatcaaatattggttatatggctcaaaatgatgctttatataatgat
ttaactggacgtgaaaatttaacgtttttcgcaagaatttatatgcgtgataaagaagat
attaaaaaacgtgtgaacctatgcagttccatggttcaattagacaatgatttagataag
aaagttgaaatgtattctggtggaatgaaacgacgcttatcttttagctattagcttttta
caaaatcctaatatccttatattagacgaacctacagttggcattgatcctaaattgcgt
cagacgatttggaaggatttaactaaagcaaaggctgaagataaatgcattatagtgaca
acgcatgtattagatgaagctacacgctgtgataagctcgtattaatgaatcaaggaaag
atattggcaacgggtacaccagatgaagtgaaaaaacaatatcacacagatacgattgaa
ggcgtatttctgaatatggagggataa

Sequence 3218
VSGHSFYRCEEEQFMIEVKNVSKSFGKQQVVDDISISFNAGEVVGLIGPSGTGKTTLIQC
ILGMEKIDGGQVTIQEHTMPNRKILSNIGYMAQNDALYNDLTGRENLTFFARIYMRDKED
IKKRVNLCSSMVQLDNDLDKKVEMYSGGMKRRLSLAISFLQNPNILILDEPTVGIDPKLR
QTIWKDLTKAKAEDKCIIVTTHVLDEATRCDKLVLMNQGKILATGTPDEVKKQYHTDTIE
GVFLNMEG*

Sequence 3219
Contig_0733_pos_9564_7870
is similar to (with p-value 7.0e-31)
 >gp:gp|AB009866|AB009866_7 Bacteriophage phi PVL proviral D
NA, complete sequence. NID: g3341907.
atggcaaagattagggattatgttacagaatatgcaaaaaagtagttaatggcgatatt
atagctagtaaaaaaaacgtgaaagcctgtcaacgccatttagatgacttgaacgattcg
gaactcccttatcattttgatgtaaagaaagctaatcacattattaagtttcttgaaatg
ttgccagatcctaaaactggtaaacaattatcgttaggcggttttcaaaaattcattgct
ggtagcttaaatggttggtacgacagacatgggtacaaaagatttacaaaagcctatata
tcaatgagcagaaaaaatggtaaaacattattgatctctggaatggcattgtacgattta
ttgatgggtaaagatccgttgaatgaacggttgattggtttgagcgccaattcaagagac
caagctggtatagcatacgatatgacattggcacaactgaaagctattagaagcgtttct
cctaaggttaaatcgatgactaagataacgccaagtgcaaaagaaatattgaatattaat
gatcgaagtaaagttaaagccgtttcaaatgaagctgcaaatttagaaggtcatcagttt
agctacgcaatcatcgatgaatatcatgaagctaaagataaaaagatttatgaaacgtta
agacgtgggcaagtgctactgcacaaccctatattaattattatctcaacagctggaact
aatttgaatggtccgatgtatgaagaatatttatatattgataagacttgacggcata
gcaaaaaatgaaaactactttgttttctgtgctgaacaagatgatgagaaagaagtatat
gacgttaaaacttggattaaatccaatccacttatggagttgccagaaatggcacaattg
ttaactaagaatattcaaccagaagttaaaactgcaattgatagtggttcaggattaaat
gggatattaataaagaatttcaatatgtggcgtgcagcaagcacagaatcttatttagat
ttcaatgattggaagaaaaatgaaatagactttgatataaatggctctaaaacttatatc
ggtttagacttatcgcgtgctgacgacttaaccgcagtatcgtttgttcatcttgatgaa
gataatcaagagtattatgtaactagtcattcgtttgtggctactaaaggtggattagat
ggcaagattgatagagactttattgattacagacaacttgcagaaagtggttattgtacg

```
attaccgatttacaaagtggaattatcaatactgaccaagttttaaattacattgagaat
tatatcgaccaatataaattagacgtacaagcgttatgttatgatccttactcaatacat
ggtgttattgcagaaattgagcgtagagattggccttatgatttagtagaaatcagacaa
gggccacaaacactatctaatccgatactggattttagactgaaagtgattaatggggac
```
5
```
atcaagcatcataaaaatccgttactagacattgcagtcaaaaatgctgtggcaaaagat
accaatgactcattaatgattgaaaagaagatgaaccgagaaaaaatagatccactcatg
gctaccatatttgcttatgttatggcttgtgaacatgaatgggacacagaaactttaatg
ccattgttcttatag
```

10   Sequence 3220
MAKIRDYVTEYAKKVVNGDIIASKKNVKACQRHLDDLNDSELPYHFDVKKANHIIKFLEM
LPDPKTGKQLSLGGFQKFIAGSLNGWYDRHGYKRFTKAYISMSRKNGKTLLISGMALYDL
LMGKDPLNERLIGLSANSRDQAGIAYDMTLAQLKAIRSVSPKVKSMTKITPSAKEILNIN
DRSKVKAVSNEAANLEGHQFSYAIIDEYHEAKDKKIYETLRRGQVLLHNPILIISTAGT
15   NLNGPMYEEYLYIDKILDGIAKNENYFVFCAEQDDEKEVYDVKTWIKSNPLMELPEMAQL
LTKNIQPEVKTAIDSGSGLNGILIKNFNMWRAASTESYLDFNDWKKNEIDFDINGSKTYI
GLDLSRADDLTAVSFVHLDEDNQEYYVTSHSFVATKGGLDGKIDRDFIDYRQLAESGYCT
ITDLQSGIINTDQVLNYIENYIDQYKLDVQALCYDPYSIHGVIAEIERRDWPYDLVEIRQ
GPQTLSNPILDFRLKVINGDIKHHKNPLLDIAVKNAVAKDTNDSLMIEKKMNREKIDPLM
20   ATIFAYVMACEHEWDTETLMPLFL*

     Sequence 3221
     Contig_0733_pos_7622_6432
     is similar to (with p-value 2.0e-16)
25   >gp:gp|AB009866|AB009866_5 Bacteriophage phi PVL proviral D
     NA, complete sequence. NID: g3341907.
```
atgaacagagatttagaacgattattgtattggcaagaacatggcacacatgcaagctat
gttggtataaacgcgctacgtaacagtgatgtatttactgctacacgtattatatctgca
gacattgcaagtaccaagttgaaagttaaaggtcacgaaacaaatacagtgatggaccaa
```
30
```
atactggatctatttaataacaatccgtattcggacttaccgggttggcactttaagttt
ataatcatcgcgaatatgctgcttaacggtcaatcttttgttgaaattgtgcgtggcaaa
aatgattttcctgttggattccacttcttacataacgacttagtaggaattgaggaaaaa
gacggcgaaattatttacaacgtaagtgaagatgtggaaggtaatgccgttaagataaca
agcgatgatatattacatttcagatatatcacattagatggatatataggatacagtccg
```
35
```
ttgtatgcactagcacatgagattggtatttctcaaggctctaagagcttcctgcgtaac
ttcttcgataatggtgggacttcgacatcagtattgaagtatagaaaagggcaaatcaat
gctgaacaattaagagaattgaaaaagaacttttcagaaagtcaattaaaaaacaacggt
ggtttagttgctatcgatgacacaatggaatttaacagactacaaattcctaccgaagta
ttgaacttcttaaatagttataagttcagcacatctcaagttgctaaagcgttcggtttg
```
40
```
ccggtatctaaactaggtattgaaacagtcaatacatctatcacacaagcaaacttagag
tatttgcaaagtacattagatccaatatttaaaatgatgattgctgaactcgaaacgaaa
atatttaaatttattgattctggtaacgaattagagtttgactcatcacgtctcattgac
attgatccagagttacaattacaacgtattactgaattgcatagtaaaggaattatttca
acagacgaagctagaagtgtatttggctatcaacctattgaacatggcgagcaaccattg
```
45
```
gttgatcttaacagagcgccacttaacactttagaaaattaccaaaaatcgaaaattgac
aaagaagtcgaaaagaactccattaaagggggtgatgagtatgacgaatag
```

     Sequence 3222
     MNRDLERLLYWQEHGTHASYVGINALRNSDVFTATRIISADIASTKLKVKGHETNTVMDQ
50   ILDLFNNNPYSDLPGWHFKFIIIANMLLNGQSFVEIVRGKNDFPVGFHFLHNDLVGIEEK
     DGEIIYNVSEDVEGNAVKITSDDILHFRYITLDGYIGYSPLYALAHEIGISQGSKSFLRN
     FFDNGGTSTSVLKYRKGQINAEQLRELKKNFSESQLKNNGGLVAIDDTMEFNRLQIPTEV
     LNFLNSYKFSTSQVAKAFGLPVSKLGIETVNTSITQANLEYLQSTLDPIFKMMIAELETK
     IFKFIDSGNELEFDSSRLIDIDPELQLQRITELHSKGIISTDEARSVFGYQPIEHGEQPL
55   VDLNRAPLNTLENYQKSKIDKEVEKNSIKGGDEYDE*

     Sequence 3223
     Contig_0733_pos_6406_5885
     is similar to (with p-value 2.0e-46)

>gp:gp|AB009866|AB009866_4 Bacteriophage phi PVL proviral D
NA, complete sequence. NID: g3341907.
atggttgtcgaaggttacgccattatctttaattcaatgagtgatgatttgggtggattt
agagaaattgtagcgcctaatgctttaaatgatgtagatgtaagtgatgtgaaatgtcta
5 atcaatcatgattttagttatgttataggacgcacacaagcaggaacgcttgagctacag
gtggatgaaaaagggctatactttaaatgccacttacctaatacatcatacgcaagagat
atttatgagaatattaaagcaggcaacgttaatcagtgcagtttctcttttacacattgcca
cctaatgactcaacggctcgtacgtggcaaaacatagataatgagtacgttcaaaccata
aataaaatcgatgaattgattgaggttagtattgttacagtgccagcctacaaagataca
10 tcggttgaagtcggtcaacgtgcgaaagacttaaagaaattcaaacagttggaacaaatg
aagatagcattggatttagaaagcctacgttttgaaacgtaa

Sequence 3224
MVVEGYAIIFNSMSDDLGGFREIVAPNALNDVDVSDVKCLINHDFSYVIGRTQAGTLELQ
15 VDEKGLYFKCHLPNTSYARDIYENIKAGNVNQCSFFYTLPPNDSTARTWQNIDNEYVQTI
NKIDELIEVSIVTVPAYKDTSVEVGQRAKDLKKFKQLEQMKIALDLESLRFET*

Sequence 3225
Contig_0733_pos_5844_4492
20 >gp:gp|X97563|BPHA3GP3_5 Bacteriophage A2 gp3 gene and 4 op
en reading frames. NID: g1523807.
atggctaatttagatgagcgcaaaaaagaaatcgctaatctgatttctaaagcgcaagaa
gcagtcgaaaaaggcgacctcgaaactgctcgtaatttaaaagctgatattgatgctcaa
aagaaagagtacgaagaactcgaacagctttcaaaagaaattgaagcgtcagcacctaaa
25 caagatgaaccacctaaagatgaaggtgcagaagttgaagataacaaagatggtaattct
ggagaagaatcagagaacaaaaccttctgatgatgaaccagaaggaacttcagatgaagaa
aaacctgatgatgcaccaaaaccagatgacaaacctgaagaaacaccagaaacacctact
attgaaaaagtagaagaaccaacagaagaagaattaaaaaaagaaaaagacaaaaaagaa
ggagcgaaacgttctatggctaaattaaaccaaatccagagacaaacgaagaaattcta
30 gcatttgaacagtacatgaaatcaaaaggggctaaacgtgacaatgttaaatctgatgac
gttggcgtaactatcccagaggatattaaatatattcctgaaaaagaagttaagacagtc
caagacttatcagaattggtacaaaaaacttcagtatcaactgcaagtgggaaatacccg
atcttaaaacgtgctaacgctaaattcaacactgttgctgaattagagaaaaaccctgag
ttagctcgtccggaattcgaaacaatcaattgggaagtagacacttatcgtggatctatt
35 ccgatttcacaagaagcattagatgattcagttgctaacttaactgctattgtttctgaa
aatattaacgaacaaaaaatcaacactttaaatgaacgtattggtgaagttttaaaagca
ttcaatcctactagtgtttctaatgttgacgacttaaaagaaattatcaacgttaaatta
gatcctggttatgaccgccaaattatctgtactcaaagtttctatcaaaaactagataca
ttaaaagatggtaacggtcgttatttactacaagacagtatcatcaacactgcaggtaac
40 actgtgttaggtatgaatgtaacagttgtgcgtgatgacttgttaggtaaaaatggagat
gcattagcatttattggtgatgtaaaacgcggtgtgttatttgcagaccgtacagacgtt
tctgttcaatggattgaaaatgaaatctacggtaaatacttaatgggtgctttccgtttc
gatgtgaaacaggctgataaaaatgctggtttcttcgtaacatttgaagagcgtttatat
tacttcatattgggcaatggatgtatacgatga
45
Sequence 3226
MANLDERKKEIANLISKAQEAVEKGDLETARNLKADIDAQKKEYEELEQLSKEIEASAPK
QDEPPKDEGAEVEDNKDGNSGEESENKPSDDEPEGTSDEEKPDDAPKPDDKPEETPETPT
IEKVEEPTEEELKKEKDKKEGAKRSMAKLNQNPETNEEILAFEQYMKSKGAKRDNVKSDD
50 VGVTIPEDIKYIPEKEVKTVQDLSELVQKTSVSTASGKYPILKRANAKFNTVAELEKNPE
LARPEFETINWEVDTYRGSIPISQEALDDSVANLTAIVSENINEQKINTLNERIGEVLKA
FNPTSVSNVDDLKEIINVKLDPGYDRQIICTQSFYQKLDTLKDGNGRYLLQDSIINTAGN
TVLGMNVTVVRDDLLGKNGDALAFIGDVKRGVLFADRTDVSVQWIENEIYGKYLMGAFRF
DVKQADKNAGFFVTFEERLYYFILGNGCIR*
55
Sequence 3227
Contig_0734_pos_2701_2249
is similar to (with p-value 9.0e-25)
>sp:sp|P00937|TRPG_YEAST ANTHRANILATE SYNTHASE COMPONENT II

(EC 4.1.3.27) (CONTAINS: GLUTAMINE AMIDOTRANSFERASE; INDOLE
-3-GLYCEROL PHOSPHATE SYNTHASE (EC 4.1.1.48) (PRAI)). >pir:p
ir|S38049|NNBY2 anthranilate synthase multifunctional enzyme
- yeast (Saccharomyces cerevisiae) >gp:gp|X75951|SC6ORF_9 S
.cerevisiae URA1, SAC1, RSD1 and TRP3 genes and 6 new orfs.
NID: g473130. >gp:gp|Z28211|SCYKL211C_1 S.cerevisiae chromos
ome XI reading frame ORF YKL211c. NID: g486376.
gtggtggtaatacatctacatttattgcattatcgtcagtatggacaagctggtgcatct
attattttattaatagtaaatattttaagtgatgaccaattaaaagaattgtattcatat
gcaacaaaccataatttagaagctctagtagaagttcatacaattagagaacttgaacgt
gcacaccaaattaaccctaaaattattggtgttaataatcgtgatttaaaacgatttgaa
accgatgttctacatacaaataaattacttaagtttaaaaagtctaattgctgctacatt
tcagagagtggcattcatacaaaagaagatgttgagaaaatagtagattcaagtattgac
ggtttacttgtaggggaggcattaatgaaaacaaatgacttaagtcagtttttgcctagt
ttaaagttaaagaagaatctctatgatagttaa

Sequence 3228
VVVIHLHLLHYRQYGQAGASIILLIVNILSDDQLKELYSYATNHNLEALVEVHTIRELER
AHQINPKIIGVNNRDLKRFETDVLHTNKLLKFKKSNCCYISESGIHTKEDVEKIVDSSID
GLLVGEALMKTNDLSQFLPSLKLKKNLYDS*

Sequence 3229
Contig_0737_pos_1135_1560
is similar to (with p-value 2.0e-19)
>gp:gp|AF001974|AF001974_1 Thermoanaerobacter ethanolicus p
utative TrkG gene, partial cds, and putative TrkA, xylose is
omerase (xylA) and xylulose kinase (xylB) genes, complete cd
s. NID: g2581794.
gtgagaccgactgtaccgaatgctgatacgacttcgaagagtattttgattaaaggtata
ttggaattgattatggtaagtataaacgtaaccataccgataaatgcaatagaaatgaga
atagttacgaaagacaactgtatgtatctttctgatatttctctattaaatatagagttg
ttttttcttttgcgtatcgtattaaaaatagcgattgttgcaataacaaaagttgtaacc
tttataccacctgcagcactcaatggtgcacctccaataaacatgagagccataagtaat
aaagctgtcggtgtttttaatgtttccaacgtcaattgtgttaaatcctgcagtccttgtt
gtcactgattggaaaaatgcatttcctattttttcaattaatcccatgtgtaacatagag
ttttga

Sequence 3230
VRPTVPNADTTSKSILIKGILELIMVSINVTIPINAIEMRIVTKDNCMYLSDISLLNIEL
FFSLRIVLKIAIVAITKVVTFIPPAALNGAPPINMRAISNKAVGVLMFPTSIVLNPAVLV
VTDWKNAFPIFSINPMCNIEF*

Sequence 3231
Contig_0737_pos_1585_2199
is similar to (with p-value 3.0e-19)
>sp:sp|P43440|NTPJ_ENTHR V-TYPE SODIUM ATP SYNTHASE SUBUNIT
J (EC 3.6.1.34) (NA(+)- TRANSLOCATING ATPASE SUBUNIT J). >g
p:gp|D17462|ENENTP_11 Enterococcus hirae ntp genes for Na+ -
ATPase subunits, complete cds. NID: g487271.
gtgcctatcaaaattaataatccagtggtagttaacacgagtttagagtgtaaggataat
tttctaaaacttttggcgttccacaaatcaaccacgactaaatgtcccaaacctcccaaa
atgataagtattggaatagtgataatgattaccggatcatttgaaaaatcgattaagttg
tttttaaaaagggcgaatcctgcgttgttaaatgcggaaactgaagtgaataaacttaaa
aatagacctttacctatgccaaattttggaataaacgataaacatagacaaagtgtacca
aataattcagtggcgatgctgtatatggctagatgtgtttaataagtttaattacaccaccg
ggttcgtcaatattccaagtaatcataaataaaattctattgttaattgatatttttctta
tttaaaaagatgagagttagcattgctacagtgacaatacctaatccaccaatttgtata
agtaataagataatgatttctccaaaaatattaaattgtgttccaacatcaactggtgat
agacctgttactgtgaatgcgctagaagctacaaacaatgcgtcaataaagttaataggt

ttcttccctgtatag

Sequence 3232
VPIKINNPVVVNTSLECKDNFLKLLAFHKSTTTKCPKPPKMISIGIVIMITGSFEKSIKL
FLKRANPALLNAETEVNKLKNRPLPMPNFGINDKHRQSVPNNSVAMLYMARCLISLITPP
GSSIFQVIINKILLLIDIFLFKKMRVSIATVTIPNPPICISNKIMISPKILNCVPTSTGD
RPVTVNALEATNNASIKLIGFFPV*

Sequence 3233
Contig_0737_pos_2249_978
is similar to (with p-value 1.0e-54)
 >gp:gp|D89592|D89592_3 Vibrio alginolyticus rhlE, KtrA and
KtrB genes, complete cds. NID: g3927863.
atgttatttctgttgacaactttaattggtgcttttctactctatttgccctatacaggg
aagaaacctattaactttattgacgcattgtttgtagcttctagcgcattcacagtaaca
ggtctatcaccagttgatgttggaacacaatttaatatttttggagaaatcattatctta
ttacttatacaaattggtggattaggtattgtcactgtagcaatgctaactctcatcttt
ttaaataagaaaatatcaattaacaatagaattttatttatgattacttggaatattgac
gaacccggtggtgtaattaaacttattaaacatctagccatatacagcatcgccactgaa
ttatttggtacactttgtctatgtttatcgtttattccaaaatttggcataggtaaaggt
ctatttttaagtttattcacttcagtttccgcatttaacaacgcaggattcgccctttt
aaaaacaacttaatcgatttttcaaatgatccggtaatcattatcactattccaatactt
atcattttgggaggtttgggacattagtcgtggttgatttgtggaacgccaaaagtttt
agaaaattatccttacactctaaactcgtgttaactaccactggattattaattttgata
ggcacggttttcttcttttttactagaaaatcaaaactctatgttacacatgggattaatt
gaaaaaataggaaatgcatttttccaatcagtgacaacaaggactgcaggatttaacaca
attgacgttggaaacattaaaacaccgacagctttattacttatggctctcatgtttatt
ggaggtgcaccattgagtgctgcaggtggtataaaggttacaacttttgttattgcaaca
atcgctatttttaatacgatacgcaaagaaaaaaacaactctatatttaatagagaaata
tcagaaagatacatacagttgtctttcgtaactattctcatttctattgcatttatcggt
atggttacgtttatacttaccataatcaattccaatatacctttaatcaaaatactcttc
gaagtcgtatcagcattcggtacagtcggtctcactatggatttaacttccgaatactat
aattggactgagtttattattatcatcgtaatgttatgtggtaaaattggattactgaat
attagtagagcgcttgttccacctaaagaccctaaaaattatagatataccaaaggacac
attcacttataa

Sequence 3234
MLFLLTTLIGAFLLYLPYTGKKPINFIDALFVASSAFTVTGLSPVDVGTQFNIFGEIIIL
LLIQIGGLGIVTVAMLTLIFLNKKISINNRILFMITWNIDEPGGVIKLIKHLAIYSIATE
LFGTLCLCLSFIPKFGIGKGLFLSLFTSVSAFNNAGFALFKNNLIDFSNDPVIIITIPIL
IILGGLGHLVVVDLWNAKSFRKLSLHSKLVLTTTGLLILIGTVFFFLLENQNSMLHMGLI
EKIGNAFFQSVTTRTAGFNTIDVGNIKTPTALLLMALMFIGGAPLSAAGGIKVTTFVIAT
IAIFNTIRKEKNNSIFNREISERYIQLSFVTILISIAFIGMVTFILTIINSNIPLIKILF
EVVSAFGTVGLTMDLTSEYYNWTEFIIIIVMLCGKIGLLNISRALVPPKDPKNYRYTKGH
IHL*

Sequence 3235
Contig_0738_pos_1259_24
 >sp:sp|O34863|UVRA_BACSU EXCINUCLEASE ABC SUBUNIT A. >gp:gp
|Z99122|BSUB0019_13 Bacillus subtilis complete genome (secti
on 19 of 21): from 3597091 to 3809700. NID: g2636029. >gp:gp
|AF017113|AF017113_13 Bacillus subtilis 300-304 degree genom
ic sequence. NID: g2618830.
atgcgtgatttaggtaatacacttattgtcgttgaacatgacgatgatactatgagagca
gctgattatttagttgatgtgggtccgggagctggtaaccacggtggagaggttgtctca
agtggtacccctaataaagtaatgaaagataaaaaatcttaactggtcaatatttaagt
ggaaaaaacgaattgaagtccctgaatacagacgagaaatcaccgatagaaagattcaa
attaaaggtgctaaaagtaataatttgaaaaatgtaaatgtagacttcccactatctgtc
ttaactgttgttacaggtgtgtcaggttctggtaaaagttcactcgtcaatgaaatttta

```
        tataaagcattagctcaaaaaattaataaatctaaagtgaaacctgggaattttgatgaa
        attaaaggaattgatcaattagataaaatcattgatattgatcaatcgccaataggtaga
        acaccacgttctaacccagccacatacactggtgtctttgatgacataagagatgtcttt
        gcacaaacgaatgaagctaaaatacgaggttatcaaaaaggtagatttagttttaatgtc
   5    aaaggtggacgatgtgaagcttgtaaaggtgatggaattataaaaattgaaatgcatttt
        ttaccagatgtctatgtaccttgtgaagtatgtgatggtaaacgctataatcgtgagact
        ttagaggtaacatacaaaggtaaaaatattgcggatgtattagaaatgactgttgaagaa
        gctacgcatttctttgaaaatattcctaagattaaacgtaaattacaaacacttgtagat
        gttgggttggggtacattactttaggtcaacaagctactacattatctggtggcgaagcg
   10   caacgtgtaaaactcgcatcagaattgcacaaacgttcaacgggcgttctatttatatt
        cttgatgaaccaactacaggattacatgtcgacgatataagtcgtttattaaaggtattg
        aatcgtatagtggaaaatggtgatacggtcgttattatcgaacacaatcttgatgttatt
        aaaacggctgatcatattattgatttaggtccagaaggcggtgaaggtggaggaacaatc
        atcgcaactggtacacctgaagagattgctcaaaataaagggtcttacactggtcaatac
   15   ttaaaaccagtattagagagagacagcgttgaatag
```

Sequence 3236
MRDLGNTLIVVEHDDDTMRAADYLVDVGPGAGNHGGEVVSSGTPNKVMKDKKSLTGQYLS
GKKRIEVPEYRREITDRKIQIKGAKSNNLKNVNVDFPLSVLTVVTGVSGSGKSSLVNEIL
YKALAQKINKSKVKPGNFDEIKGIDQLDKIIDIDQSPIGRTPRSNPATYTGVFDDIRDVF
AQTNEAKIRGYQKGRFSFNVKGGRCEACKGDGIIKIEMHFLPDVYVPCEVCDGKRYNRET
LEVTYKGKNIADVLEMTVEEATHFFENIPKIKRKLQTLVDVGLGYITLGQQATTLSGGEA
QRVKLASELHKRSTGRSIYILDEPTTGLHVDDISRLLKVLNRIVENGDTVVIIEHNLDVI
KTADHIIDLGPEGGEGGGTIIATGTPEEIAQNKGSYTGQYLKPVLERDSVE*

Sequence 3237
Contig_0741_pos_5345_6325
is similar to (with p-value 1.0e-21)
 >sp:sp|P14940|ADH_ALCEU ALCOHOL DEHYDROGENASE (EC 1.1.1.1).
 >pir:pir|A30196|A30196 alcohol dehydrogenase (EC 1.1.1.1) -
 Alcaligenes eutrophus >gp:gp|J03362|AFAADH_1 A.eutrophus al
 cohol dehydrogenase (ADH) gene, complete cds. NID: g141899.

```
        atgtttaaaaagattgctactataataggttcgacattatttggtacagttttattcgca
        aaagtgaaagaaaagcgtagttataaaagttttttacaagagaaatgattagaatatca
        ggaatgaaaaagacatttgaaagtatagatgacgcgaaaaaagctttgaatgagactaaa
   35   tatcaaactgcaggtaaatataatggaacaacatatgaatttaagcataaagttcaaata
        agagattattatggttcttagtctatgttgttaatgatcatggtcttccagatcaacgc
        acggtcttatatgtacatggaggcgcatggttccaagatcctttggaaaatcattttgaa
        tatttagacttactcgttgatgcgctcgatgctagggtgattatgcccgtatatcctaaa
        ataccacacagagattatcgtacgacatttgaattattaacaaaaatatataagcgatta
   40   ttgactaaaattgatgaacctgaaaacttggtcatcattggagattcagccgggggacaa
        attgcattagcttttgcacaaatgttaaaaaaagagcaactcagtcaacctggccatatt
        gttcttatttcaccggtgcttgatgcgacatttaagaatccagaagcaagaaaatatgaa
        aaagaagatccaatgcttggaattgaaggcagtaaatatcttgtagagttatgggctggt
   45   gatgcaccactagatgactataagatgtctccaatgaatggtgatttagaaggcctagga
        catattacacttactgtaggaaccaaagaaacattatatcctgatgcagttaagttctct
        cacatgttaaatgataaaggaataaagcatcagtttatcccaggttacaatttatttcat
        atttatcccttattccctatcccagagcgtcaacgctttttagaacagcttaaaaaaatc
        attgtcacaaaagagttataa
```

Sequence 3238
MFKKIATIIGSTLFGTVLFAKVKEKRSYKSFLQEKMIRISGMKKTFESIDDAKKALNETK
YQTAGKYNGTTYEFKHKVQIRDYYGSLVYVVNDHGLPDQRTVLYVHGGAWFQDPLENHFE
YLDLLVDALDARVIMPVYPKIPHRDYRTTFELLTKIYKRLLTKIDEPENLVIIGDSAGGQ
IALAFAQMLKKEQLSQPGHIVLISPVLDATFKNPEARKYEKEDPMLGIEGSKYLVELWAG
DAPLDDYKMSPMNGDLEGLGHITLTVGTKETLYPDAVKFSHMLNDKGIKHQFIPGYNLFH
IYPLFPIPERQRFLEQLKKIIVTKEL*

Sequence 3239

Contig_0741_pos_3244_2630
is similar to (with p-value 2.0e-21)
  >gp:gp|L38252|ACCEST_2 Acinetobacter lwoffii orf1 and ester
ase (est) genes, complete cds. NID: g1209221.
5 atgttatcagatatacttccaacaggttatgaaattggtgtttttaaaaggtaaagttaaa
cctggctgtacagtagccattgtaggtgctggtcctgtaggtttagcagcattacttaca
gcacaattctattcaccttcaaaaattattatgattgatttagatgataatagattagaa
accgctaaagaactaggtgctacgcatttaattaactctaaagagactgaaaccgcaatt
aaaaaggtaaaatcgttaaatccacgtggtgttgatgttgctattgaagctgtcggaatt
10 ccacaaacctttgatttatgtcaaaatttaattggtgtcgatggtacgattgctaatgtt
ggtgtgcatgggttacctgtacaacttgatatagataaattatggattaaaaatattaac
gtaactactggtttagtttcaggaaatacaactgaagaattacttgaagcgttaaaaagc
aaaataatacaaccagaacaactcgtgacacattatagtaaactgagtgaaatcgaaagt
gcctatgatttatttagaaatgcaacagatcataaagcgattaaattaatcatagagaat
15 gatatcacaatttaa

Sequence 3240
MLSDILPTGYEIGVLKGKVKPGCTVAIVGAGPVGLAALLTAQFYSPSKIIMIDLDDNRLE
TAKELGATHLINSKETETAIKKVKSLNPRGVDVAIEAVGIPQTFDLCQNLIGVDGTIANV
20 GVHGLPVQLDIDKLWIKNINVTTGLVSGNTTEELLEALKSKIIQPEQLVTHYSKLSEIES
AYDLFRNATDHKAIKLIIENDITI*

Sequence 3241
Contig_0743_pos_0_688
25 is similar to (with p-value 1.0e-50)
  >sp:sp|P70814|RIBG_BACAM RIBOFLAVIN-SPECIFIC DEAMINASE (EC
3.5.4.-). >gp:gp|X95955|BARIBGENS_1 B.amyloliquefaciens ribB
, ribG, ribA, ribH & ribT genes. NID: g1592687.
atggatgatgctattcaactagcaaaaatggtaaatggacaaacaggtgttaatccacca
30 gtaggatccgttgttgttaaaaacggtaggattgtaggtttaggtgcacatttaaaaaag
ggagataaacatgccgaagtacaagctattgaaatggcaggtttaaatacccaaggtgct
accatatacgtttcattagaaccttgcacacaccatggttcaacaccaccttgtgtgcat
aaaatcattgaagcgggcatatctaaggtcatctatgctgttaaagatactactttagta
agtaagggtgacgagattctgagagaagctggtatagaggttgaatttcaatataatgaa
35 aatgcagctgcattataccgtgactttttttactgctaaaagaaacgaagttccagaagta
actgtaaaggtctcatctagtctagatggtaaacaagcaacagactttaatgaaagtaag
tggataacaaacaaagaagttaaagaagatgtttatcaattaagacatgagcatgatgca
gttattactgggcgtagaaccattgaagcagacaatccattgtatacaaccagggttcct
gatggaaagcatccgattcgagttattctttctaagaaaggtcaactcgattttaatcaa
40 caaatatttaaagatactgcatcCCCTG

Sequence 3242
MDDAIQLAKMVNGQTGVNPPVGSVVVKNGRIVGLGAHLKKGDKHAEVQAIEMAGLNTQGA
TIYVSLEPCTHHGSTPPCVHKIIEAGISKVIYAVKDTTLVSKGDEILREAGIEVEFQYNE
45 NAAALYRDFFTAKRNEVPEVTVKVSSSLDGKQATDFNESKWITNKEVKEDVYQLRHEHDA
VITGRRTIEADNPLYTTRVPDGKHPIRVILSKKGQLDFNQQIFKDTASPX

Sequence 3243
Contig_0744_pos_2445_3218
50 is similar to (with p-value 5.0e-69)
  >gp:gp|U96107|SCU96107_3 Staphylococcus carnosus N5,N10-met
hylenetetrahydromethanopterin reductase homolog, SceB precur
sor (sceB) and putative transmembrane protein genes, complet
e cds, and putative Na+/H+ antiporter NhaC (nhaC) gene, part
55 ial cds. NID: g2735503.
atgaaaaaaatcgctacagctacaattgcaactgcaggaatcgctactttcgcatttgca
caccatgacgcacaagcagcagaacaaaataatgatgggtacaatccaaacgacccttat
tcatatagctacacttacacaatcgatgctgaaggtaactaccactacacttggaaaggt
aactggagtccagatcgtgtaaatacttcatataactataataattataataactacaac

tactatggttacaataactatagcaactacaataactacagtaattacaacaattacaac
aactatcaatcaaacaacacgcaatcacaaagaacaactcaaccgactggtggtttaggc
gcaagctattcaacatcaagtagtaatgttcacgttacaacaacttctgcgccatcatca
aacggtgtatctttatcaaacgctcgctcagcatctggtaacttatacacttcaggtcaa
5   tgtacatattatgtatttgacagagtaggtggcaaaatcggttcaacgtggggtaacgca
aacaactgggcaaacgctgcagcacgttctggttacacagtaaacaattcgcctgctaaa
ggtgcaatcttacaaacgtcacaaggtgcatacggacacgtagcatacgttgaaggtgta
aacagcaatggttcaatcagagtttcagaaatgaactacggtcacggtgcaggtgttgtc
acttcacgtacaatctctgcgagccaagctgcttcatataactatattcactaa
10

Sequence 3244
MKKIATATIATAGIATFAFAHHDAQAAEQNNDGYNPNDPYSYSYTYTIDAEGNYHYTWKG
NWSPDRVNTSYNYNNYNYYGYNNYSNYNNYSNYNNYNNYQSNNTQSQRTTQPTGGLG
ASYSTSSSNVHVTTTSAPSSNGVSLSNARSASGNLYTSGQCTYYVFDRVGGKIGSTWGNA
15   NNWANAAARSGYTVNNSPAKGAILQTSQGAYGHVAYVEGVNSNGSIRVSEMNYGHGAGVV
TSRTISASQAASYNYIH*

Sequence 3245
Contig_0744_pos_7402_7740
20   is similar to (with p-value 9.0e-30)
        >sp:sp|P37279|ATCS_SYNP7 CATION-TRANSPORTING ATPASE PACS (E
        C 3.6.1.-). >pir:pir|S36741|S36741 cation-transporting ATPas
        e pacS - Synechococcus sp. >gp:gp|D16437|SYOPACS_1 Synechoco
        ccus sp. DNA for PacS, complete cds. NID: g435124.
25   atggtagaaccaatcactgaatcgccactacttttttcaacaggtatagactcaccagtt
agcatggattcgtcaatagaagtatcacctttagtgactttgccatctacaggtatcttt
tcgccgggtttattagtaaagtatctccgactttaacttatcaagtggaagcataatt
tctttattttctttaattactcgtgcttctttcgcttgtaaatttaacaattcgcttaat
gcattggtagtctgtgatttggcacgtgcttctaaatatttaccaagaagaattaacgta
30   attaaaatagcacttgtttcaaaatataaatgcgggtga

Sequence 3246
MVEPITESPLLFSTGIDSPVSMDSSIEVSPLVTLPSTGIFSPGFISKVSPTLTLSSGSII
SLFSLITRASFACKFNNSLNALVVCDLARASKYLPRRINVIKIALVSKYKCG*
35

Sequence 3247
Contig_0744_pos_9912_8887
        is similar to (with p-value 3.0e-66)
        >sp:sp|P56157|DP3A_HELPY DNA POLYMERASE III, ALPHA CHAIN (E
40   C 2.7.7.7). >gp:gp|AE000646|HPAE000646_3 Helicobacter pylori
        section 124 of 134 of the complete genome. NID: g2314635.
        atgggatttgatgaaacgagtttaaatgagatttcaaaacttattccacataaattaggt
ataactcttgaagaagcataccaaaagccagagtttaaagcatttgttcatcgtaatcat
agaaatgaacgttggtttgaagtgagtaaaaagttagagggattaccaagacatacgtct
45   acgcatgctgcaggtatcattatcaatgatcaaccattattcaaatttgccccattaaca
actggtgatacaggattattaacgcagtggactatgacagaagcggaacgtataggatta
ttaaaaattgatttcttgggattacgcaatctatcaattattcatcaaattattttacaa
gttaaaaaggatttaaatataaatattgatatagaagctataccttatgatgataaaaaa
gtttttgatttattatcaaacggtgacactacaggtatatttcaattggaatcagacggt
50   gttagaagcgtattaaaaagattgcaacccgaacattttgaagatatcgtagctgtcaca
tcattatatagaccaggaccaatggaagaaataccaacttatataacccgtagacataat
cctaaccaatttgcttatttacatccagatttagaaccaatcttaaaaaacacatatggt
gttatcatttatcaagaacaaataatgctaatagcaagtcaagttgctggtttttagttat
ggtgaagcagatattttaagaagggcaatgagtaaaaagaatcgtgcaatcttagaaagt
55   gagcgtcaacatttcattgatggtgcaaaaaataacggttacgatgaacagataagtaag
caaattttgatttaatacttaagtttgcagattatgggttcccacgtgcccatgctgtt
agttactcaaaaattgcatacattatgagctatttaaaagtgccacaaatcagtacagaa
ccgttatgcctggcttcaaatatgaagcaccaccaaatcaaaataaaatcaatccttatg
aattaa

813

Sequence 3248
MGFDETSLNEISKLIPHKLGITLEEAYQKPEFKAFVHRNHRNERWFEVSKKLEGLPRHTS
THAAGIIINDQPLFKFAPLTTGDTGLLTQWTMTEAERIGLLKIDFLGLRNLSIIHQIILQ
VKKDLNINIDIEAIPYDDKKVFDLLSNGDTTGIFQLESDGVRSVLKRLQPEHFEDIVAVT
SLYRPGPMEEIPTYITRRHNPNQFAYLHPDLEPILKNTYGVIIYQEQIMLIASQVAGFSY
GEADILRRAMSKKNRAILESERQHFIDGAKNNGYDEQISKQIFDLILKFADYGFPRAHAV
SYSKIAYIMSYLKVPQISTEPLCLASNMKHHQIKIKSILMN*

Sequence 3249
Contig_0746_pos_2558_3991
is similar to (with p-value 2.0e-45)
 >sp:sp|P26648|SUFI_ECOLI SUFI PROTEIN PRECURSOR. >gp:gp|U28
377|ECU28377_116 Escherichia coli K-12 genome; approximately
 65 to 68 minutes. NID: g882431. >gp:gp|AE000384|AE000384_1
Escherichia coli K-12 MG1655 section 274 of 400 of the compl
ete genome. NID: g2367186.
atgtataataaagtttttgcaattttaattataatttttccataataattattgcgtct
aatgatactttcgcagaaagtaagaatgatatgatgaatatgaaagaagataagaaaaat
acaatggatatgacaaatatgaaacatcatgacgaaagaaagaaattaaattcttcacaa
ggaaaaaatgaaataatatttcctaaagttgcagagtcaaaaaaagataacaatggttat
aaaaattatacattaaaagctcaggaaggaaagacagagttttacaaaaataattttct
aatactctaggctacaatggaaatttacttggaccaactttaaaattaaaaaaaggagat
aaagttaaaattaagttaataaataacttagatgaaaatacaacatttcattggcatgga
ttagaagtaaatggaaaagtggatggagggccttctcaggttataaaaccaggaaaagaa
aaaactataaaatttgaggttaatcaagattctgctacgttatggtatcacccccacccc
tctccaaatacagctaaacaagtttataatggcttatcaggattattatatatagaagat
agtaaaaagaataattatcctagtgattatggaaaaaatgatttgcctataataatccaa
gataaaacatttgtatctaaaaaattaaattattcaaaaacgaaagacgaagatggcact
caaggtgatactgttcttgtgaacggaatagtaaaccccaaactgacaacaaaagaagag
aaaatacgtttgagactttaaatggttctaatgctcgagatttaaatcttaagctaagt
aataatcaaagttttgagtatattgcttcagatggcggtcaattaaaaaaacgctaaaaaa
ttaaaagaaattaatttagctccttcagaaagaaaagaaatagtaatagatttatctaaa
atgaaaggcgagaaaatcagtctggttgataatgataaaactgtaattttaccgattagt
aacaaagagaaaagttctaacaaaggtaatacaccaaaagtaagtaaaaaaataaaatta
gaaggtatgaatgataatgttaccattaatggtaataaattcgatcctaaaagaatagat
tttacacaaaagttaaaccagaaagaagtatgggaaattgaaacgtcaaagataaaatg
ggtggtatgaaacatcctttccacatccatggaacgcaatttaaagtttttatctgtggat
ggggagaaaccaccaaaagatatgaggggtaaaaaagatgttatatctttggaacctgga
caaaaagctaaaatagaggttgtatttaaaaatactggaacatacatgtttcactgtcat
atacttgagcatgaagataatggaatgatgggtcaaataaaagtaacaaactaa

Sequence 3250
MYNKVFAILIIIFSIIIIASNDTFAESKNDMMNMKEDKKNTMDMTNMKHHDERKKLNSSQ
GKNEIIFPKVAESKKDNNGYKNYTLKAQEGKTEFYKNNFSNTLGYNGNLLGPTLKLKKGD
KVKIKLINNLDENTTFHWHGLEVNGKVDGGPSQVIKPGKEKTIKFEVNQDSATLWYHPHP
SPNTAKQVYNGLSGLLYIEDSKKNNYPSDYGKNDLPIIIQDKTFVSKKLNYSKTKDEDGT
QGDTVLVNGIVNPKLTTKEEKIRLRLLNGSNARDLNLKLSNNQSFEYIASDGGQLKNAKK
LKEINLAPSERKEIVIDLSKMKGEKISLVDNDKTVILPISNKEKSSNKGNTPKVSKKIKL
EGMNDNVTINGNKFDPKRIDFTQKLNQKEVWEIENVKDKMGGMKHPFHIHGTQFKVLSVD
GEKPPKDMRGKKDVISLEPGQKAKIEVVFKNTGTYMFHCHILEHEDNGMMGQIKVTN*

Sequence 3251
Contig_0748_pos_3358_4230
is similar to (with p-value 1.0e-38)
 >sp:sp|P44948|FPG_HAEIN FORMAMIDOPYRIMIDINE-DNA GLYCOSYLASE
 (EC 3.2.2.23) (FAPY-DNA GLYCOSYLASE). >pir:pir|A64104|A6410
4 formamidopyrimidine-DNA glycosylase (fpg) homolog - Haemop
hilus influenzae (strain Rd KW20) >gp:gp|U32776|U32776_1 Hae

814

mophilus influenzae Rd section 91 of 163 of the complete gen
ome. NID: g1573969.
atgcctgaactacctgaagttgaacatgttaaaagaggtattgagccatttataaaaagt
gcaaaaatagagaaagtaacttttgctaaaaatgtaattaacggtaagaataataaccgt

5 gagacgattataaaaggtatggaattagatacttttaaaaaacttactgaaggttatgtt
ataaaaaaagttgaaagaagaagtaagtacattattttttatatagcggatcatgacgat
gatagaatcttagttagtcatttaggtatggcaggcggattctttgttgttaataacctt
gatgagataagtacaccgaattatcgaaagcattggcaagtcattttcgatttggataat
aaacaaaaattagtctattctgatatcagacggtttggagaaattagaaatatagtcaat

10 tttgatagttatccatctttattagaaatcgctccagaacccatttgaagaggtggcattt
gaacactatttagaatgtttgacaatgaaaaaatataagaataaaccaataaaacaaacg
attcttgatcatcgtgttatagcaggagctggaaatatctatgcctgtgaagctttattc
agagctggtattactccggataaaattactaattcactcactaaacaagaaagaaaatcc
ctcttttattatgttcgagaagttttagaagagggtataaaatatggaggtactagtatt

15 tcagattataggcatgcagatggtaaaactggacaaatgcaattacatttaaatgtatat
aaacaaaaaagtgcaaggtttgtggtcattcgattgaaacaaaagtgatagctggtaga
aatagtcattttgcccaaactgtcagagataa

Sequence 3252
20 MPELPEVEHVKRGIEPFIKSAKIEKVTFAKNVINGKNNNRETIIKGMELDTFKKLTEGYV
IKKVERRSKYIIFYIADHDDDRILVSHLGMAGGFFVVNNLDEISTPNYRKHWQVIFDLDN
KQKLVYSDIRRFGEIRNIVNFDSYPSLLEIAPEPFEEVAFEHYLECLTMKKYKNKPIKQT
ILDHRVIAGAGNIYACEALFRAGITPDKITNSLTKQERKSLFYYVREVLEEGIKYGGTSI
SDYRHADGKTGQMQLHLNVYKQKKCKVCGHSIETKVIAGRNSHFCPNCQR*

25

Sequence 3253
Contig_0748_pos_4883_5752
is similar to (with p-value 4.0e-89)
  >gp:gp|Z82038|CTZ82038_4 C.thermosaccharolyticum etfB, etfA
30 , hbd, thlA and actA genes. NID: g1667352. >gp:gp|Z92974|TTB
CSOPRN_6 T.thermosaccharolyticum BCS operon DNA. NID: g19033
26.
gtgtttggtggtgtatttaaggatatacctgcctatgaactaggtgcaacagttattcgt
caaattttagaacatagtcaaatagatcctaatgaaatcaatgaagttattctaggaaac
35 gtattacaggcaggtcaaggacaaaatcctgctcgtattgctgcgattcatggtggtgtg
ccagaagcggtaccttctttttactgtaaataaagtttgcggttctggattaaaagcgatt
caacttgcctatcaatctattgtagcgggagataatgagattgttatcgctggaggcatg
gaaagtatgtctcaatctccaatgcttcttaaaaatagtcgtttcggttttaaaatggga
aatcaaactttagaagatagtatgatagctgatggtttaactgataagtttaatgattac
40 catatgggtatcacagccgaaaatctagttgaacagtatcagattagtcgtaaagaacaa
gatcaatttgcattcgattctcaacaaaaagcatcacgtgcacaacaagctggtgtattt
gatgctgaaattgtacctgtagaggtaccacaacgtaaaggcgaccccctaattattct
caagatgaaggcattagacctcaaacgacaattgataagttagcacaactccgtccagca
tttaaaaaagatggatcagtaactgctggtaatgcatccggtatcaatgacggtgctgct
45 gctatgctcgttatgacggaggacaaagcgaaagcattgggcttacaacctatagctgta
ttagatagtttttggtgcgagtggtgtggcgccttcaattatgggtattcgacgcacaaaa
acaaaccaaagaagaaatagcaactgttaa

Sequence 3254
50 VFGGVFKDIPAYELGATVIRQILEHSQIDPNEINEVILGNVLQAGQGQNPARIAAIHGGV
PEAVPSFTVNKVCGSGLKAIQLAYQSIVAGDNEIVIAGGMESMSQSPMLLKNSRFGFKMG
NQTLEDSMIADGLTDKFNDYHMGITAENLVEQYQISRKEQDQFAFDSQQKASRAQQAGVF
DAEIVPVEVPQRKGDPLIISQDEGIRPQTTIDKLAQLRPAFKKDGSVTAGNASGINDGAA
AMLVMTEDKAKALGLQPIAVLDSFGASGVAPSIMGIRRTKTNQRRNSNC*

55

Sequence 3255
Contig_0753_pos_1053_58
is similar to (with p-value 3.0e-40)
  >gp:gp|AB011003|AB011003_1 Candida albicans CaUAP1 gene for

UDP-N-acetylglucosamine pyrophosphorylase, complete cds. NI
D: g3273313.
atgttagacaaaaatcaattggaaaagtataaccaagagcatttgtatgaatatgaaaaa
ttaatgagtagtaatgaaaagaatgctttagatgaaaaagtagatcagttaaatcttgca
gaaattcaagatttatatcaagatttatatgttaatagaaaaactattgatgatgtatct
tctgtatctgaagtcaaatatgaagtgaaatcacgactcaatgaagaagaacgacataca
tatgaacaaaaaggttatgaggcaatacgaaatggtgaatttgctgtattattgatggct
ggaggacaaggtacgcgtttaggatataaagggcctaaaggttcttttgaaatagagggt
acgagtttatttgaacttcaggcgcgtcaactgattcgtttaaaagaagaaaccggccac
acaattaattggtatattatgacaagtgacattaatcataaagatacaatagagtatttt
aaacaacataaatattttaactatgatgccaatcatattcatttctttaagcaagataac
attgttgctttaagtgaagaaggaaagcttgtttttaaatagagatggacatataatggaa
acacctaatggtaatgggggtgtattcaagtctcttaagaaagcaggataccttgataag
atgcaacaagatcacgtcaaatatatcttcttaaataacattgataatgtcttagttaaa
gttttagacccgttatttgccggtttttacagtgacacaaagtaaagacatcacatcaaaa
acaattcaacctaaagatagtgaaagtgtaggtcggcttgtaaatgttgattgtaaagac
actgtgttagagtattctgaattaatgtcatatttccaggagctgcactccagttcattt
ctgaaactaaaatacttccatcgctattcacgcgctctacaaacgctacgtgaccatagt
aaccagcgtcagtttgtgcaattgagcctactgtag

Sequence 3256
MLDKNQLEKYNQEHLYEYEKLMSSNEKNALDEKVDQLNLAEIQDLYQDLYVNRKTIDDVS
SVSEVKYEVKSRLNEEERHTYEQKGYEAIRNGEFAVLLMAGGQGTRLGYKGPKGSFEIEG
TSLFELQARQLIRLKEETGHTINWYIMTSDINHKDTIEYFKQHKYFNYDANHIHFFKQDN
IVALSEEGKLVLNRDGHIMETPNGNGGVFKSLKKAGYLDKMQQDHVKYIFLNNIDNVLVK
VLDPLFAGFTVTQSKDITSKTIQPKDSESVGRLVNVDCKDTVLEYSELMSYFQELHSSSF
LKLKYFHRYSRALQTLRDHSNQRQFVQLSLL*

Sequence 3257
Contig_0754_pos_7866_0
is similar to (with p-value 4.0e-44)
 >sp:sp|P36649|YACK_ECOLI PROBABLE 53.4 KD BLUE-COPPER PROTE
IN YACQ PRECURSOR. >gp:gp|AE000121|AE000121_8 Escherichia co
li K-12 MG1655 section 11 of 400 of the complete genome. NID
: g1786306.
atgtataataaagtttttgcaatttttaattataattttttccataataattattgcgtct
aatgatactttcgcagaaagtaagaatgatatgatgaatatgaaagaagataagaaaaat
acaatggatatgacaaatatgaaacatcatgacgaaagaaagaaattaaattcttcacaa
ggaaaaaatgaaataatatttcctaaagttgcagagtcaaaaaaagataacaatggttat
aaaaattatacattaaaagctcaggaaggaaagacagagtttttacaaaaataattttttct
aatactctaggctacaatggaaatttacttggaccaactttaaaattaaaaaaaggagat
aaagttaaaattaagttaataaataacttagatgaaaatacaacatttcattggcatgga
ttagaagtaaatggaaaagtggatggagggccttctcaggttataaaaccaggaaaagaa
aaaactataaaatttgaggttaatcaagattctgctacgttatggtatcacccccacccc
tctccaaatacagctaaacaagtttataatggcttatcaggattattatatatagaagat
agtaaaaagaataattatcctagtgattatggaaaaaatgatttgcctataataatccaa
gataaacatttgtatctcaaaaaattaaattattcaaaaacgaaagacgaagatggcact
caaggtgatactgttcttgtgaacggaatagtaaaccccaaactgacaacaaaagaagag
aaaatacgtttgagactttaaatggttctaatgctcgagatttaaatcttaagctaagt
aataatcaaagtttttgagtatattgcttcagatggcggtcaattaaaaaaacgctaaaaaa
ttaaaagaaattaatttagctccttcagaaagaaaagaaatagtaatagatttatctaaa
atgaaaggcgagaaaatcagtctggttgataatgataaaactgtaattttaccgattagt
aacaaagagaaaagttctaacaaaggtaatacaccaaaagtaagtaaaaaaataaaatta
gaaggtatgaatgataatgttaccattaatggtaataaattcgatcctaaaagaatagat
tttacacaaaagttaaaccagaaagaagtatgggaaattgaaaacgtcaaagataaaatg
ggtggtatgaaacatcctttccacatccat

Sequence 3258
MYNKVFAILIIIFSIIIIASNDTFAESKNDMMNMKEDKKNTMDMTNMKHHDERKKLNSSQ

GKNEIIFPKVAESKKDNNGYKNYTLKAQEGKTEFYKNNFSNTLGYNGNLLGPTLKLKKGD
KVKIKLINNLDENTTFHWHGLEVNGKVDGGPSQVIKPGKEKTIKFEVNQDSATLWYHPHP
SPNTAKQVYNGLSGLLYIEDSKKNNYPSDYGKNDLPIIIQDKTFVSKKLNYSKTKDEDGT
QGDTVLVNGIVNPKLTTKEEKIRLRLLNGSNARDLNLKLSNNQSFEYIASDGGQLKNAKK
LKEINLAPSERKEIVIDLSKMKGEKISLVDNDKTVILPISNKEKSSNKGNTPKVSKKIKL
EGMNDNVTINGNKFDPKRIDFTQKLNQKEVWEIENVKDKMGGMKHPFHIH

Sequence 3259
Contig_0755_pos_2032_3393
is similar to (with p-value 8.0e-68)
  >gp:gp|L41217|RICNRAMP_1 Oryza Sativa integral membrane pro
tein (OsNramp) mRNA. NID: g2231131. >gp:gp|S81897|S81897_1 O
sNramp1=Nramp1 homolog/Bcg product homolog [Oryza sativa, in
dica, cv. IR 36, etiolated shoots, mRNA, 1967 nt]. NID: g147
0319.
atgggggtgattttattgaattccaataataacaatcatgaacaacaacgaagtttagat
gaaatcaataacaccataaacttcaatcataatgatagtgcaagtcaaaaatttctggct
tttttaggaccgggattgcttgttgcagttggttacatggatcctggaaattggattaca
tccatgcaaggaggagcacaatatggctataccttgttattcataatcttaatctctagc
ttatctgctatgctgttacaaagtatgactgtgagattaggaatagcaactggtatggat
ttagcacaaatgacacgtcattttttaaataagcctgtagcaattatgttctggattatt
gcagaattagcaattatcgctactgatattgcagaagtatataggtagcgctatcgcatta
gatttaatcttcggcataccattaattgtaggcgcattaatcactgtatttgatgtatttt
ttattattattcatcatgaaatttggctttagaaagattgaagctatcgtgggaacgtta
atctttaccgtattggccattttgtatttgaagtttatatttcttctccacaaataaca
gatatgcttaatggttttgtgcctcataaagaaattattacaaaccaagggatactttat
attgcactaggtatcataggtgctactattatgccacataacttatatttacattcttct
attgtacaatctcgaaaatatgatagacacagtattcatgaaaaagcacaagcgattaag
tatgctactatcgactctaatatacagctatccatcgcttttgtagtcaattgcttatta
cttacacttggtgcagcgctatttttggaactaaaactgaagatttgggtggtttttat
gatctttatttggctctaaaaacagaacctgctttaggtgcaacgcttggcggtattatg
agtactttatttgctgttgccctttttagcttctggtcaaaattcaactataacgggaacg
ttagcaggccaaattgtgatggaaggattttcttaaattatccattccaaattggttacgt
cgtcttatcactcggtctttagcagtgatacctgttatcatttgtcttatagtatttaaa
ggaaatactgaaaaaattgaacaattacttgtcttttctcaagtgttcttgagtattgct
ttgccatttttcgttaataccgcttcaattagctacaagtaatcaaaatcttatgggtcct
tttaagaataaaacatggattaacatcatttcttggttactcataattgtcttaagtgga
cttaacgtatatcttatcattcaaacattccaagaattatga

Sequence 3260
MGVILLNSNNNNHEQQRSLDEINNTINFNHNDSASQKFLAFLGPGLLVAVGYMDPGNWIT
SMQGGAQYGYTLLFIILISSLSAMLLQSMTVRLGIATGMDLAQMTRHFLNKPVAIMFWII
AELAIIATDIAEVIGSAIALDLIFGIPLIVGALITVFDVFLLLFIMKFGFRKIEAIVGTL
IFTVLAIFVFEVYISSPQITDMLNGFVPHKEIITNQGILYIALGIIGATIMPHNLYLHSS
IVQSRKYDRHSIHEKAQAIKYATIDSNIQLSIAFVVNCLLLTLGAALFFGTKTEDLGGFY
DLYLALKTEPALGATLGGIMSTLFAVALLASGQNSTITGTLAGQIVMEGFLKLSIPNWLR
RLITRSLAVIPVIICLIVFKGNTEKIEQLLVFSQVFLSIALPFSLIPLQLATSNQNLMGP
FKNKTWINIISWLLIIVLSGLNVYLIIQTFQEL*

Sequence 3261
Contig_0755_pos_6203_5673
 No hits found
atgggattaaataaattagggtatagtcttaatgacaaaaatccgacccatattcatcaa
gcagaaaaagatttgcataatttagcacctcaagttagagggatagtaggcgacgaaatt
actatgatgcttcaacaaaacgaaggacatgttgcagtagtttggagtggcgttgctgca
ccacttgtacaggaaaatactcgttataattacgtgatacctaaagaaggctctaaccta
tggtttgataatatggtgatacctaaaactgcacaaaataaagaaggtgcgtataagttt
atgaatttcttactagacgcgcaaaatagtgcccagaatacggaatgggtaggatatgca
acacctaataaagcagctcgaagtaagttgcctaaaaaggtaagaaatgattatagattt

tatccatcaaatcaagaacagcaacggttagaagtctataaagatttaggtcaaacgtct
ctcagtgaatataatgaaagctttttaaattttaaaatgtctttaaaatag

Sequence 3262

MGLNKLGYSLNDKNPTHIHQAEKDLHNLAPQVRGIVGDEITMMLQQNEGHVAVVWSGVAA
PLVQENTRYNYVIPKEGSNLWFDNMVIPKTAQNKEGAYKFMNFLLDAQNSAQNTEWVGYA
TPNKAARSKLPKKVRNDYRFYPSNQEQQRLEVYKDLGQTSLSEYNESFLNFKMSLK*

Sequence 3263

Contig_0755_pos_3989_3612
is similar to (with p-value 8.0e-25)
　>sp:sp|P23861|POTD_ECOLI SPERMIDINE/PUTRESCINE-BINDING PERI
PLASMIC PROTEIN PRECURSOR (SPBP). >pir:pir|D40840|D40840 spe
rmidine/putrescine transport protein D - Escherichia coli >g
p:gp|D90747|D90747_2 Escherichia coli genomic DNA.(25.1 - 25
.5 min). NID: g1651548. >gp:gp|AE000212|AE000212_9 Escherich
ia coli K-12 MG1655 section 102 of 400 of the complete genom
e. NID: g1787358. >gp:gp|M64519|ECOPOTABCD_4 E.coli transpor
t protein (potA, potB, potC and potD) genes, complete cds. N
ID: g147325.
atgtcatctgaaatgtttttagtttctcttattgcgtcattagtcattttaggtttttca
acgctattaggttttgttggtacaatggtaattgaaggtagaaaaaaccttgctgctagt
ttactcatagcagcggctatcgtaggttttatttacgactaatttaatcgcaatggtttta
tggatgattgctgcgattagacttttttgcaaaaaaagataaaacagatgtaaatgaaaat
gctacggctcaacttcgtcaaaaccattcaaagagccaaagtgattggaatcatcaacaa
aaccaacaacagaaagatgcttgggatcctgaacaagaaatcaacaaacaaaaaaaggac
gatccatatatatattaa

Sequence 3264

MSSEMFLVSLIASLVILGFSTLLGFVGTMVIEGRKNLAASLLIAAAIVGLFTTNLIAMVL
WMIAAIRLFAKKDKTDVNENATAQLRQNHSKSQSDWNHQQNQQQKDAWDPEQEINKQKKD
DPYIY*

Sequence 3265

Contig_0756_pos_5439_6113
is similar to (with p-value 3.0e-39)
　>gp:gp|U67196|TMU67196_1 Thermotoga maritima DNA-binding re
sponse regulator (drrA) and histidine protein kinase (hpkA)
genes, complete cds, thymidne/pyrimidine phosphorylase homol
og gene, partial cds. complete cds. NID: g1575576.
atgattaattgcttaatcgtagacgatgataaaaagttattgcaatatgtttcaagtcat
ttagaaagagaaagtattcaaacacatactttcacaagtggagaagcatcactagatttt
cttgaaaataaaaatgttgatattgcgatagtagatattatgatgagtggaatggatggt
tttgagctttgtcagactttgaaagatgattatcatattcctgtcataatgttaacagct
agagatgcattaagtgataaagaacgtgcatttctaagtgggactgacgattatgtcact
aaaccttttgaggttaaagaattattatttagaattaaagctgtcttaagacgatatcaa
attaatgctgataacgagttacaacttggcaacttaatattaaatcagtcttacatggaa
attactgtgggttcaaaaacgatgaatcttccaaacaaagaatttcagttgttattttta
ttagcctctaatcctaaacatattttcactcgagatgatattattgaaaaaatttggggc
ttcgattatgaaggagatgatcgtactgttgatgttcatattaaaagattacgtcaacgt
ttatctaaattgaaatcatcagtatcaattcaaactgtaagaggtcaaggatatagggtg
gaccaaaatgtttaa

Sequence 3266

MINCLIVDDDKKLLQYVSSHLERESIQTHTFTSGEASLDFLENKNVDIAIVDIMMSGMDG
FELCQTLKDDYHIPVIMLTARDALSDKERAFLSGTDDYVTKPFEVKELLFRIKAVLRRYQ
INADNELQLGNLILNQSYMEITVGSKTMNLPNKEFQLLFLLASNPKHIFTRDDIIEKIWG
FDYEGDDRTVDVHIKRLRQRLSKLKSSVSIQTVRGQGYRVDQNV*

Sequence 3267
Contig_0757_pos_1852_2325
is similar to (with p-value 1.0e-26)
>gp:gp|AJ000084|PMAJ84_3 Proteus mirabilis ccm and pat gene
s and partial ygbA gene. NID: g3395515.
atgaatgctaatgcccaagtacgtagatacttccctagtttacttagttatcgtcgttct
gagatagacatggaaaatatggatagcataatacaagagagtggcctaggtttatttgct
gtagagttgaaagagacgggagaatggttagggtttataggtgtgaattacgtttcaaaa
gatagccattaccctttaaagagttaccgtttttatgaaataggttggaggctaattcca
gaagtatggggaaatggtctagctacagaaggggcagaagccgtaatgaaatatgctaga
gataaaggaattaaagaattatatagtttttacttctgaaaataatttgccttctagaaa
gtcatggaaaaattaggaatgactttttttagacaattttgaatatccgaatcttagtaaa
taccatccccttaaacgtcatgtaagatattataaagagctacttccttcttga

Sequence 3268
MNANAQVRRYFPSLLSYRRSEIDMENMDSIIQESGLGLFAVELKETGEWLGFIGVNYVSK
DSHYPFKELPFYEIGWRLIPEVWGNGLATEGAEAVMKYARDKGIKELYSFTSENNLPSRK
VMEKLGMTFLDNFEYPNLSKYHPLKRHVRYYKELLPS*

Sequence 3269
Contig_0757_pos_4021_3656
is similar to (with p-value 2.0e-29)
>sp:sp|P46378|FAS6_RHOFA HYPOTHETICAL 21.1 KD PROTEIN IN FA
SCIATION LOCUS (ORF6). >pir:pir|F55578|F55578 hypothetical p
rotein 2 (ipt 3' region) - Rhodococcus fascians plasmid pFiD
188 >gp:gp|Z29635|RFCCIPTFD_6 R.fascians (D188) genes for P4
50 cytochrome, isopentenyltransferase and ferredoxine. NID:
g455000.
atggacaaaaagcttcaaaagaatattgagaaacgtcataaagaagaacaaaaacaacgt
gaagctaatcagaaacaacgtattaaagacatgaaaaaaactcaaaaatacgaagagcaa
gttggcttaactcctggtaaaatagatcacgaaattgagaaaaaaggcgaaaaactagaa
aaagataatcgtaaagatattaaaaaattagataaaaagcttcaaaagaatattgaaaaa
cgtcataaagaagaacaaaaacagcgtgaagaagcagagaaagctagaaaaaaagaattt
aaaaaatatgaaaattacgtggctgacagtgtcgtaaaacaacataaggaatcaaatcat
tcttaa

Sequence 3270
MDKKLQKNIEKRHKEEQKQREANQKQRIKDMKKTQKYEEQVGLTPGKIDHEIEKKGEKLE
KDNRKDIKKLDKKLQKNIEKRHKEEQKQREEAEKARKKEFKKYENYVADSVVKQHKESNH
S*

Sequence 3271
Contig_0757_pos_3193_2627
 No hits found
atgaaaagaattgctgtttattgtggtgcaagtaaagggaaaaacccatcttatgttaaa
gaggcatacgaattaggaaaatatatggctgaacaaggatacgagcttgtattcggagca
ggatcagtcggcattatgggagctattcaagatggcatacttgagcatggcggtaaagct
atcggtgtcatgcctaaaatgttagatgaacgagaaataacaagccaaaaagtaagtgag
cttatattagtagattctatgcatgaacgaaaaaataaaatgactgaacttgccgatgct
tttattatggctccaggcggtgctggttcattagaagaattttttgaaatgtatagttgg
gctcaaattggtatacaccaaaagcctataggtgtatttaatttaaatggattctttgag
ccactacaacacttaatcgaccatatgattaaagaaggatttattgatgagaaatatcaa
aagcttgcacctttatatgatactaaagaatcactcatcgaaggacttaaacattacaaa
ccacttggtgtacgtacatacgattaa

Sequence 3272
MKRIAVYCGASKGKNPSYVKEAYELGKYMAEQGYELVFGAGSVGIMGAIQDGILEHGGKA
IGVMPKMLDEREITSQKVSELILVDSMHERKNKMTELADAFIMAPGGAGSLEEFFEMYSW
AQIGIHQKPIGVFNLNGFFEPLQHLIDHMIKEGFIDEKYQKLAPLYDTKESLIEGLKHYK

PLGVRTYD*

Sequence 3273
Contig_0759_pos_3773_3006
is similar to (with p-value 9.0e-58)
  >gp:gp|U38892|SSU38892_2 Synechocystis sp. ruvB gene, compl
ete cds, and secA gene, partial cds. NID: g1256587.
atgtttaaaataggaaatttagaattacaatctcgtttacttttaggtactggaaaattt
gaaaatgaagaggttcagtcaaaagcaattgaggcatctgaaacaaatgtacttacattt
gcagtaagacgtatgaatttatatgatcgtaacctacctaacccacttgcaaacgttaat
ttaaaagattttatcacttttccaaatactgcaggtgccaaaacagctcaagaagctatc
agaattgctgaaattgctaatcacgcaggtgtatgtgacatgattaaagtcgaagtcatt
ggtgatgacgaaacattattacctgatccattcgaaacatacgaggcatgcaaagtattg
ttagaaaaaggttacactgtttgtccttacatctctaacgatttagtttagctcaacgt
ttagaagaattgggtgtacacgcagttatgccacttgcatcccctattggtacaggaaga
ggtattaataacccattaaatttaagttatattatcgaaaatgctagtgtacctgtaatc
gtagatgctggtattggttcccctaaagatgcgtgtcatgccatggagcttggcgcagat
ggtattttactcaacacagccatttcagcggcaaaagatcctgtgaaaatggctgaagca
atgaaattaggtataaatgctggcagactttcatatgaagctggacgcattcctgttaag
tatactgcacaagcatctagtccatcagaaggtttagggttcttgtaa

Sequence 3274
MFKIGNLELQSRLLLGTGKFENEEVQSKAIEASETNVLTFAVRRMNLYDRNLPNPLANVN
LKDFITFPNTAGAKTAQEAIRIAEIANHAGVCDMIKVEVIGDDETLLPDPFETYEACKVL
LEKGYTVCPYISNDLVLAQRLEELGVHAVMPLASPIGTGRGINNPLNLSYIIENASVPVI
VDAGIGSPKDACHAMELGADGILLNTAISAAKDPVKMAEAMKLGINAGRLSYEAGRIPVK
YTAQASSPSEGLGFL*

Sequence 3275
Contig_0763_pos_1317_2216
is similar to (with p-value 1.0e-47)
  >gp:gp|U76260|PAU76260_2 Peptostreptococcus asaccharolyticu
s alpha- and beta-subunits of L-serine dehydratase (sdhB) an
d (sdhA) genes, complete cds. NID: g2315864.
atgtttgattcaattagagagacaatagattattcagttgaaaataacatcagttttgct
gatatgatgattaatgatgaaatggaaagagaaggtaaatctcgcgaagaagtgcgtgat
ttaatgagacaaaacttaaatgttatgcgtgaggcagttgaaaaaggtacgactggcgat
ggtgtggaaagtgtaacaggttatacaggtcatgatgctgctaaacttcgagattacaac
gagaataatcatgcattatcaggtcatgaaatgattgatgcagtcaaaggtgcagttgca
acgaacgaagtcaatgcagcaatgggtattatttgtgctactccgacagctggttcctcg
ggaacgattcccggcgtaatatttaaattagaaaaaactcataatatcactgaagatcaa
atgatagattttctattcacatcagctttattcgggcgagtagttgcaaacaacgcgagc
gttgccggtgcaactggtggttgtcaagccgaagtgggttcggcatctgcaatggctgca
gctgctgcagtatcaatttttaacgggtcaccagaacaatcaggacatgccatggcattg
gcaattagtaacttattaggcttagtttgcgatccagttgctggtttagttgaaattcca
tgtgtaatgcgtaatgctattggttcaggaaatgcattaatatctgcagaccttgcacta
gctggagttgaaagtcaaattccagttgatgaagtcataggtgctatggatagagtaggt
cgtaatttacctgcatcattaagagaaacaggtttaggcggtttagcaggtacacctact
ggcgaagaattaaacgtaaaatattcggcgaagcagacaacatggttaaaaataaataa

Sequence 3276
MFDSIRETIDYSVENNISFADMMINDEMEREGKSREEVRDLMRQNLNVMREAVEKGTTGD
GVESVTGYTGHDAAKLRDYNENNHALSGHEMIDAVKGAVATNEVNAAMGIICATPTAGSS
GTIPGVIFKLEKTHNITEDQMIDFLFTSALFGRVVANNASVAGATGGCQAEVGSASAMAA
AAAVSIFNGSPEQSGHAMALAISNLLGLVCDPVAGLVEIPCVMRNAIGSGNALISADLAL
AGVESQIPVDEVIGAMDRVGRNLPASLRETGLGGLAGTPTGEEIKRKIFGEADNMVKNK*

Sequence 3277

Contig_0765_pos_2844_3362
is similar to (with p-value 3.0e-17)
 >sp:sp|P04928|SANT_PLAFN S-ANTIGEN PROTEIN PRECURSOR. >pir:
 pir|B22011|YAZQN7 S-antigen precursor - Plasmodium falciparu
 m (strain NF7/Ghana) >gp:gp|M10130|PFASA7_1 Plasdmodium falc
 iparum (isolate NF7) S antigen gene, complete cds. NID: g160
 670.
gtgcttctgcattgcctccttcttctgccttcgttggctcagattgaggtgcttctgcat
tgcttccttcttctgtcttcgttggctcagattgagctgcttctgcattgcctccttctt
ctgccttcgttggctcagattgagctgcttttgcgttgcttccttcttctgtcttcgttg
gctcagattgaggtgcttctgcattgcttccttcttctgccttcgttggctcagattgag
ctgcttttacgttgcttccttcttctgtcttcgttggctcagattgagctgcttctgcat
tgcctccttcttctgccttcgttggctcagattgagctgcttctgcattgcctccttctt
ctgccttcgttggctcagattgaggtgcttctgcattgcctccttcttctgccttcgttg
gctcagattgagctgcttctgcattgcctccttcttctgccttcgttggctcagattgag
gtgcttctgcattgcctccttcttctgcttgcaatgtag

Sequence 3278
VLLHCLLLLPSLAQIEVLLHCFLLLSSLAQIELLLHCLLLLPSLAQIELLLRCFLLLSSL
AQIEVLLHCFLLLPSLAQIELLLRCFLLLSSLAQIELLLHCLLLLPSLAQIELLLHCLLL
LPSLAQIEVLLHCLLLLPSLAQIELLLHCLLLLPSLAQIEVLLHCLLLLLAM*

Sequence 3279
Contig_0768_pos_1456_1800
is similar to (with p-value 4.0e-59)
 >pir:pir|I67760|I67760 transposase (insertion sequence IS10
) - Escherichia coli >gp:gp|S67119|S67119_2 BST=somatotropin
 ...BST/beta-Gal fusion protein [Escherichia coli, LBB84, pla
 smid pXT107, IS10L/R-1, PlasmidTransposonInsertionMutant, 3
 genes, 1679 nt]. NID: g455674.
atgcagattgaagaaaccttccgagacttgaaaagtcctgcctacggactaggcctacgc
catagccgaacgagcagctcagagcgtttttgatatcatgctgctaatcgccctgatgctt
caactaacatgttggcttgcgggcgttcatgctcagaaacaaggttgggacaagcacttc
caggctaacacagtcagaaatcgaaacgtactctcaacagttcgcttaggcatggaagtt
ttgcggcattctggctacacaataacaagggaagactcactcgtggctgcaaccctgctt
actcaaaatctattcacacatggttacgttttggggaaattatga

Sequence 3280
MQIEETFRDLKSPAYGLGLRHSRTSSSERFDIMLLIALMLQLTCWLAGVHAQKQGWDKHF
QANTVRNRNVLSTVRLGMEVLRHSGYTITREDSLVAATLLTQNLFTHGYVLGKL*

Sequence 3281
Contig_0769_pos_8606_8947
is similar to (with p-value 6.0e-18)
 >sp:sp|Q41364|SOT1_SPIOL 2-OXOGLUTARATE/MALATE TRANSLOCATOR
 PRECURSOR. >gp:gp|A47930|A47930_1 Sequence 1 from Patent WO
 9534654. NID: g2301793. >gp:gp|U13238|SOU13238_1 Spinacia ol
 eracea envelope membrane 2-oxoglutarate/malate translocator
 (SODiT1) mRNA, chloroplast mRNA encoding chloroplast protein
 , complete cds. NID: g595680.
atgagtttgcttaaccatgggataaagcctaacttgtttaattgttctgccattaataca
agaactgagaaccaaacgagtgtattccatgcgcctgtttcatttaaaatatctgaccac
gctaatacacctgttaatagtaacaatgctaaagcaataaatgcagtgagcgtggcatca
acattaatgaagcttcctaatacccacaaagccaatgctatgataaagacaccaaccatc
aatttttcggctatagacatatgtcccatttcttctagttgttcagtagcccatttttta
gcgttaggcgtttctttaacagtaggtggggtataatttataa

Sequence 3282
MSLLNHGIKPNLFNCSAINTRTENQTSVFHAPVSFKISDHANTPVNSNNAKAINAVSVAS

TLMKLPNTHKANAMIKTPTINFSAIDICPISSSCSVAHFLALGVSLTVGGYNL*

Sequence 3283
Contig_0769_pos_9387_8242
>sp:sp|Q41364|SOT1_SPIOL 2-OXOGLUTARATE/MALATE TRANSLOCATOR
PRECURSOR. >gp:gp|A47930|A47930_1 Sequence 1 from Patent WO
9534654. NID: g2301793. >gp:gp|U13238|SOU13238_1 Spinacia ol
eracea envelope membrane 2-oxoglutarate/malate translocator
(SODiT1) mRNA, chloroplast mRNA encoding chloroplast protein
, complete cds. NID: g595680.
atggcctttttcatttcaagaggatttgtaaaaacagggctaggtcgacgtattgctctg
caattcgttaaattatttggaaagaaaacgcttggtttggcttattcacttgttggtgtt
gaccttatcttagctcctgctacgccaagtaatacagcacgtgctggtggtattatgttt
ccaatcattaagtccttgtcagagtcatttggttcatcgccgagagatggttctgagaga
aaaatgggtgcgtttttaatctttactgagttccaaggtaatttaattacttcagctatg
tttttaacagctatggccggtaaccctatagcgcaaagtttagctgaaaaaacggcacac
gttcaaattacatggatgaattggtttgttgctgctattatacccggattgatttctctc
atcgttgtccctttcattatttataaattatacccacctactgttaaagaaacgcctaac
gctaaaaatgggctactgaacaactagaagaaatgggacatatgtctatagccgaaaaa
ttgatggttggtgtctttatcatagcattggctttgtgggtattaggaagcttcattaat
gttgatgccacgctcactgcatttattgctttagcattgttactattaacaggtgtatta
gcgtggtcagatatttttaaatgaaacaggcgcatggaatacactcgtttggttctcagtt
cttgtattaatggcagaacaattaaacaagttaggctttatcccatggttaagcaaactc
attgctcaaggtttgaatggctttagttggcctatcgtttttagttttactcatcttgttt
tatttctactcacattatttattcgcaagtgcaacagcacatgtcagcgccatgtacgcc
gcgttactcggtgttgcagtcgcttcggtgcaccgccattattcagtgcattaatgtta
gggttctttggtaacttactggcatcaacaacacactatagtagtggaccagcgcctata
ttatacgcagatggctatgttacacaaaagcgctggtggactatgaatattgtacttggt
atagtctattttattatttggattggtgtaggttcactatggatgaaactcattggtatg
atgtaa

Sequence 3284
MAFFISRGFVKTGLGRRIALQFVKLFGKKTLGLAYSLVGVDLILAPATPSNTARAGGIMF
PIIKSLSESFGSSPRDGSERKMGAFLIFTEFQGNLITSAMFLTAMAGNPIAQSLAEKTAH
VQITWMNWFVAAIIPGLISLIVVPFIIYKLYPPTVKETPNAKKWATEQLEEMGHMSIAEK
LMVGVFIIALALWVLGSFINVDATLTAFIALALLLLTGVLAWSDILNETGAWNTLVWFSV
LVLMAEQLNKLGFIPWLSKLIAQGLNGFSWPIVLVLLILFYFYSHYLFASATAHVSAMYA
ALLGVAVASGAPPLFSALMLGFFGNLLASTTHYSSGPAPILYADGYVTQKRWWTMNIVLG
IVYFIIWIGVGSLWMKLIGMM*

Sequence 3285
Contig_0771_pos_5868_0
is similar to (with p-value 1.0e-34)
>sp:sp|P45476|YHCC_ECOLI HYPOTHETICAL 34.6 KD PROTEIN IN AR
CB-GLTB INTERGENIC REGION (F309). >gp:gp|U18997|ECOUW67_140
Escherichia coli K-12 chromosomal region from 67.4 to 76.0 m
inutes. NID: g606010.
atgcatgaaaaatggagtgaaggtcaatatattgcatactttcaggcgtttacaaatacg
catgcacctgttgaagtactaaaagaaaaatatgaacctgtcttaaaagaagatggcgtc
gttggtttatcaatcgcgacaagacctgattgtttgcctgatgatgttgtagaatattta
gctgaacttaatcagcgcacttacttatgggtagaattgggcctacaaactgtgcatcag
tcaacttccgatttaataaatcgtgctcatgatatgcaaacatactatgacggcgtaaca
aaattacgcaaacataatatataatgtttgtacgcacataatcaacggcttaccaggtgaa
aattatgacatgatgatggagactgctaaagaagtcgctcagatggacgttcaaggtatt
aaaattcatttattacacttgctaaaaggaacgcctatggttaaacaatatgaaaaaggt
atgctcgagtttatg

Sequence 3286
MHEKWSEGQYIAYFQAFTNTHAPVEVLKEKYEPVLKEDGVVGLSIATRPDCLPDDVVEYL

AELNQRTYLWVELGLQTVHQSTSDLINRAHDMQTYYDGVTKLRKHNINVCTHIINGLPGE
NYDMMMETAKEVAQMDVQGIKIHLLHLLKGTPMVKQYEKGMLEFM

Sequence 3287
Contig_0774_pos_1335_622
is similar to (with p-value 2.0e-89)
  >sp:sp|P39149|UPP_BACSU URACIL PHOSPHORIBOSYLTRANSFERASE (E
C 2.4.2.9) (UMP PYROPHOSPHORYLASE) (UPRTASE). >pir:pir|S4936
4|S49364 uracil phosphoribosyltransferase - Bacillus subtili
s >gp:gp|Z38002|BSSPORUPP_10 B.subtilis spoII-R, glyC and up
p genes. NID: g556877. >gp:gp|Z99122|BSUB0019_186 Bacillus s
ubtilis complete genome (section 19 of 21): from 3597091 to
3809700. NID: g2636029.
atgaagctaaagagagagttcacgcactcacatctaaatatccattatataattaataga
catacacattggaggaaaatgattatgagtaaagtacatgtttttgatcacccattaata
caacacaaactaagttatattagagatgctcgcactggaacaaaagagtttagggaactt
gtagatgaagtcggtatgttaatggcttatgaagtaactagagacttagaactgcaagat
gttgaaatacaaacacctgtgactaaaatgacagctaaacgtttggcgggtaaaaagtta
gcaattgtacctattttaagagctggtctaggcatgacagatggtgtgttaagtcttgtt
cctgctgctagggtaggacatataggactatatagagatccagagactcttgaagcggta
gagtactttgcgaaaatgcctcaagacatcgatgaacgtcaaattattgtggttgatcct
atgcttgctactggtgcttcagctattgaagcaatttcttcattaaaaaaacgtggagct
aaaagtatacgtttttatgtgtttaatagctgcccctgaaggcgttgaaaaaatgcaagaa
gcacacccagatgtagatatatatattgcggcattagatgaaaaattaaatgacaaagcg
tatattacaccaggtttaggtgatgcaggggatagattattcggtactaaataa

Sequence 3288
MKLKREFTHSHLNIHYIINRHTHWRKMIMSKVHVFDHPLIQHKLSYIRDARTGTKEFREL
VDEVGMLMAYEVTRDLELQDVEIQTPVTKMTAKRLAGKKLAIVPILRAGLGMTDGVLSLV
PAARVGHIGLYRDPETLEAVEYFAKMPQDIDERQIIVVDPMLATGASAIEAISSLKKRGA
KSIRFMCLIAAPEGVEKMQEAHPDVDIYIAALDEKLNDKAYITPGLGDAGDRLFGTK*

Sequence 3289
Contig_0775_pos_5036_4413
is similar to (with p-value 2.0e-43)
  >sp:sp|P44310|KGUA_HAEIN GUANYLATE KINASE (EC 2.7.4.8) (GMP
  KINASE). >pir:pir|H64139|H64139 5'guanylate kinase (gmk) ho
molog - Haemophilus influenzae (strain Rd KW20) >gp:gp|U3284
8|U32848_2 Haemophilus influenzae Rd section 163 of 163 of t
he complete genome. NID: g3212240.
atggataaggaaaaggactgttaattgttctttcaggcccttcaggtgttggaaaggga
actgttagaaagaagatatttgaagacccaactacttcatataagtattctatatcaatg
acgacacgtcatatgcgtgaaggtgaaattgatggtgtagattacttcttaaaacaaag
gaagaatttgaggcgttaattaaagacgaccagtttattgagtatgcacaatatgtaggt
aattactatggtacacctgtacaatatgtaaaggatactatggaagaaggtcatgacgtc
ttttagaaatcgaagttgaaggtgctaagcaagtaagaaagaaatttccagatgcgttg
ttcatattttttagcgcctccaagtttagatgacttgaaagaacgtcttgttggtagagga
actgaatcagatgaaaagattcaaagtcgtgtgaacgaggcacgaaaagaagtagaaatg
atgaatttatacgactacgttgtagttaacgacgaggttgaactcgctaagaatcgaatt
cagtcaatagttgaagctgagcatttaaaaagagagcgaatcgaagctaaatatagaaaa
atgttactggaggtcaaaaaataa

Sequence 3290
MDKEKGLLIVLSGPSGVGKGTVRKKIFEDPTTSYKYSISMTTRHMREGEIDGVDYFFKTK
EEFEALIKDDQFIEYAQYVGNYYGTPVQYVKDTMEEGHDVFLEIEVEGAKQVRKKFPDAL
FIFLAPPSLDDLKERLVGRGTESDEKIQSRVNEARKEVEMMNLYDYVVVNDEVELAKNRI
QSIVEAEHLKRERIEAKYRKMLLEVKK*

Sequence 3291

Contig_0779_pos_0_1180
is similar to (with p-value 6.0e-61)
  >gp:gp|U31756|BSU31756_2 Bacillus subtilis gamma-aminobutyr
ate permease (gabP) gene and spore coat protein (cotA) gene,
5   partial cds. NID: g969025.
atgattgccattggggggatgtataggaactggtctattcatgacttctggtggagctata
catgacgcaggtgcattgggtgctttgattgcctatgcagttattggagcgatggtgttc
tttctaatgacgtcgttaggggagatggcgacatatttgcctgtgtcaggttcatttagt
acttatgctacacgctttgtcgatccttcactaggtttttgctttaggatggaattattgg
10   ttcaactgggtgataacagtagcagctgatgttactattgcagcgcaagttatacaatat
tggtcccctatgcaaggtataccagcttgggtctggagttgtattttccttattattatt
ttcgcgcttaattctttatccgttagagtatatggagagagtgaatattggttcgcactt
atcaaagtagttacagtcatcatatttataggaattggtatcttaactattttagggatt
atgggtggagaatttgtaggatttgatacgtttacaaaaggagatgggccaatactaggt
15   gggaatttaggaggtagcttgctatcaattcttggtgtatttctagtcgcaggcttctca
ttccaaggaactgaacttattggtattacagcaggtgaatctgaaaatccagaaagagca
gttccaaaagcgattaaacaagtattttggcgtatacttttattttacattctagctatt
ttcattattggtatgttgattccatatgatagtaaggcattaatgggcggtggtgatagt
atagctacttcacctttttacattagtatttaagaatgctggattagcttttgctgcttca
20   tttatgaatgctgttatattaacaagtgtattatcagcaggtaactcaggaatgtatgct
tcaacaagaatgttatattcgatgagtaaagataaattagcttataattcttttggaaaa
acaaataaaagtggcgtaccttatgtatctctaattgcaactggagtactagtcattctt
attttcgcattgcaacatttaagtggagatgcatatgaatacattgtagctgctagcgga
atgactggtttttattgcttgggttggtatagcaatcagtcactttagatttagacgcgca
25   tttgataaacaaaattatgataaatcaaaattaaaatatC

Sequence 3292
MIAIGGCIGTGLFMTSGGAIHDAGALGALIAYAVIGAMVFFLMTSLGEMATYLPVSGSFS
TYATRFVDPSLGFALGWNYWFNWVITVAADVTIAAQVIQYWSPMQGIPAWVWSCIFLIII
30   FALNSLSVRVYGESEYWFALIKVVTVIIFIGIGILTILGIMGGEFVGFDTFTKGDGPILG
GNLGGSLLSILGVFLVAGFSFQGTELIGITAGESENPERAVPKAIKQVFWRILLFYILAI
FIIGMLIPYDSKALMGGGDSIATSPFTLVFKNAGLAFAASFMNAVILTSVLSAGNSGMYA
STRMLYSMSKDKLAYNSFGKTNKSGVPYVSLIATGVLVILIFALQHLSGDAYEYIVAASG
MTGFIAWVGIAISHFRFRRAFDKQNYDKSKLKYX
35

Sequence 3293
Contig_0786_pos_2712_3827
is similar to (with p-value 7.0e-97)
  >sp:sp|P96612|DDL_BACSU D-ALANINE--D-ALANINE LIGASE (EC 6.3
40   .2.4) (D-ALANYLALANINE SYNTHETASE). >gp:gp|AB001488|AB001488
_40 Bacillus subtilis genome sequence, 148 kb sequence of th
e region between 35 and 47 degree. NID: g1881226. >gp:gp|Z99
106|BSUB0003_103 Bacillus subtilis complete genome (section
3 of 21): from 402751 to 611850. NID: g2632653.
45   atggaggaaaacgaaatgacaaaagaaaatatttgtatcgtttttggaggtaaaagtgca
gaacacgatgtttcaattttaactgcacaaaatgttttaaacgcaattgataaagaacga
tatcaagttgatatcatttatataacaaacgatggtgaatggaagaaaaaagataatatt
acacaagaaataaaaaatactgatgaactcgtcattaacgatgtagagactggagaaatc
tcacagttactcagtaaaggtagtttaggaaaatcatatgatgcagtattcccattattg
50   catggtccaaatggagaagatggaactatccaaggtcttttttgaagtacttgatatacca
tatgtaggtaatggtgtgttagctgcttcaagctcaatggataaactcgtgatgaaacaa
ttatttgagcatagaggtttacctcaattaccttatattagcttttttaagaagtgagtat
gaaaaatatgaaaataatatcattaaattagttaatgataagttaacatatccggtatttt
gtaaaacctgctaatctcggttcaagtgttggtataagtaaatgtaacaatgaagaagaa
55   ttaaaatctgggatagctgaagcattccaatttgatcgtaaacttgtcattgaacaaggg
attaatgctagagagatcgaagtagctgtcttaggtaacgattatcctgaaacgacatgg
cctggtgaggttgttaaggatgtagcgtttttatgattataaatcaaagtataaagacggt
aagattagattagatattccagcagatttagatcaagatgttcaaatgacattaagaaac
atggcattagaggcctttaaagctactgattgttcgggattagttcgtgcagattttcttt

```
gttactgatgataatcaaatttatattaatgaaacaaatgccatgccaggatttactgca
tatagtatgtatccaaacttatggaaaaatatgggcttatcttaccctgatttaattgct
aaattgattgatttagctaaagaacgttatgaagataaaaagaaaaataaatattcacaa
atgtttcatctaaaaccgatcgttttttatatttaa
```

5

Sequence 3294
MEENEMTKENICIVFGGKSAEHDVSILTAQNVLNAIDKERYQVDIIYITNDGEWKKKDNI
TQEIKNTDELVINDVETGEISQLLSKGSLGKSYDAVFPLLHGPNGEDGTIQGLFEVLDIP
YVGNGVLAASSSMDKLVMKQLFEHRGLPQLPYISFLRSEYEKYENNIIKLVNDKLTYPVF

10
VKPANLGSSVGISKCNNEEELKSGIAEAFQFDRKLVIEQGINAREIEVAVLGNDYPETTW
PGEVVKDVAFYDYKSKYKDGKIRLDIPADLDQDVQMTLRNMALEAFKATDCSGLVRADFF
VTDDNQIYINETNAMPGFTAYSMYPNLWKNMGLSYPDLIAKLIDLAKERYEDKKKNKYSQ
MFHLKPIVFYI*

15
Sequence 3295
Contig_0788_pos_6565_7542
is similar to (with p-value 8.0e-33)
 >gp:gp|AB011003|AB011003_1 Candida albicans CaUAP1 gene for
 UDP-N-acetylglucosamine pyrophosphorylase, complete cds. NI

20
D: g3273313.
atgaaaattttttgattatgaagatatacaattaattcccaataaatgtattgttgaaagc
agatctgagtgtaatacttcggttaaatttgggcctcgtacttttaaattgccagttgtt
ccagcaaatatgcaaacagtcatgaatgaagaacttgcacaatggtttgcagaaaacgat
tatttttatatcatgcatagatttaatgaagaaaatagaattccatttataaaaaaaatg

25
catcatgcagggttatttgcttctatttctgttggagttaaagaaaacgaatttaatttt
attgaaaaattagcttcttcatcgctcataccagaatatataacaattgatattgctcat
ggtcactcaaattcagttataaatatgattaagcatataaaaaaacatttaccaaatagt
tttgtgatagctggtaatgttgggacgcctgaaggagtaagagaacttgagaatgccggt
gcagatgctacaaaagtaggtattggtccaggaagagtatgtattactaaaattaaaact

30
ggatttggtacaggaggttggcaactttctgcgttaaatctttgtaataaggcagctaga
aaacctattattgcagatggagggttaagaacccacggtgatatagccaaatcaattcgt
tttggtgccactatggtaatgattggctctttatttgctgcccacgaggaatcaccgggt
gaaaccgtcgagctagatggcaaaaaatataaagaatattttggtagtgcctcagaatat
caaaaaggtgaacataagaacgttgaaggtaaaaaaatgtttgtagaacacaaaggatct

35
cttaaagatacccttactgaaatggaacaagatttacagagttcaatttcatatgcagga
ggaaaggacttgaagtcattaagaacagttgattacgtcatagtaagaaattcaatctttt
aatggtgatagagattag
```

Sequence 3296
40
MKIFDYEDIQLIPNKCIVESRSECNTSVKFGPRTFKLPVVPANMQTVMNEELAQWFAEND
YFYIMHRFNEENRIPFIKKMHHAGLFASISVGVKENEFNFIEKLASSSLIPEYITIDIAH
GHSNSVINMIKHIKKHLPNSFVIAGNVGTPEGVRELENAGADATKVGIGPGRVCITKIKT
GFGTGGWQLSALNLCNKAARKPIIADGGLRTHGDIAKSIRFGATMVMIGSLFAAHEESPG
ETVELDGKKYKEYFGSASEYQKGEHKNVEGKKMFVEHKGSLKDTLTEMEQDLQSSISYAG

45
GKDLKSLRTVDYVIVRNSIFNGDRD*

Sequence 3297
Contig_0788_pos_4151_3402
is similar to (with p-value 6.0e-40)
50
 >sp:sp|P27442|GUAC_ASCSU GMP REDUCTASE (EC 1.6.6.8) (GUANOS
 INE 5'-MONOPHOSPHATE OXIDOREDUCTASE). >gp:gp|M82838|NEMGMP_1
 Ascaris lumbricoides GMP reductase mRNA, complete cds. NID:
 g159660.
atgacaagtgacattaatcataaagatacaatagagtattttaaacaacataaatattt

55
aactatgatgccaatcatattcatttctttaagcaagataacattgttgctttaagtgaa
gaaggaaagcttgtttttaaatagagatggacatataatggaaacacctaatggtaatggg
ggtgtattcaagtctcttaagaaagcaggatacctttgataagatgcaacaagatcacgtc
aaatatatcttcttaaataacattgataatgtcttagttaaagttttagacccgttatt
gccggttttacagtgacacaaagtaaagacatcacatcaaaaacaattcaacctaaagat

agtgaaagtgtaggtcggcttgtaaatgttgattgtaaagacactgtgttagagtattct
gaattagatacagacatagcaaatcaatttaacaatgctaatataggtatccatgctttt
aaactaggtttcattaccagtgctgttgatagagaattaccgtatcatttggcaatcaag
caattaaagcaattagatgaaaattttggtgtggttgaacgtcctacattaaagtttgaa
5   ttattttattttgatatatttaggtacggtacaagctttgtaacattacaagttccaaga
gaagaagaatttttcaccttttaaaaaataaagaaggtaaggatagtgtacatacagctaca
gaagatttaaaacgaatggatttgatttaa


Sequence 3298
10  MTSDINHKDTIEYFKQHKYFNYDANHIHFFKQDNIVALSEEGKLVLNRDGHIMETPNGNG
GVFKSLKKAGYLDKMQQDHVKYIFLNNIDNVLVKVLDPLFAGFTVTQSKDITSKTIQPKD
SESVGRLVNVDCKDTVLEYSELDTDIANQFNNANIGIHAFKLGFITSAVDRELPYHLAIK
QLKQLDENFGVVERPTLKFELFYFDIFRYGTSFVTLQVPREEEFSPLKNKEGKDSVHTAT
EDLKRMDLI*
15

Sequence 3299
Contig_0795_pos_2866_4128
>gp:gp|AJ224946|CGA224946_1 Corynebacterium glutamicum DNA
for L-Malate:quinone oxidoreductase. NID: g3059092.
20  atgcacatgagtgaagcaaatcataaaaacatcgttgttgtaggtgcaggaattattggt
acgtcagtagcgacaatgctttcaaaagtaagtcctaactggcatatcgatatgtttgaa
agactagaaggcgctggtattgaaagttcaaatgaaaataataatgctgggacaggtcat
gcggcattatgtgaattaaactatacagttgaacaagatgatggttcaattgatgcatct
aaagcgcaagaaattaatgaacaattcgaattatctagacaattctggggtaatttagtt
25  aaaaatggtgatatttctaatcctgaagaatttattcaaccattacctcatatcagtttc
gttatgggaccaacaaacgttaacttttttaagaaaacgttatgaaacactaagaactctt
ccaatgttcgatacaatcgaatatacagaagacatggaaacaatgagaaaatggatgcca
ttaatgatggaaaatcgtgaaccaggtcatcaaatggcagcaagtaaaattgatgaaggt
acagatgtgaactatggtgcgttaacacgtaagttagcacattacttagaacaaaaatct
30  aatgtttcattaaaatacaatcatgatgttgtagatttaacacaaagagaagatggcaaa
tgggaagttgtcgttgaaaatagagaaactaaagaaaaagtaactaaaatagcagataaa
gtgtttattggtgctggcggtcactctattccgttattacaaaaatctggcgttaaacaa
agagaacacctaggtggtttcccaatcagtggtcaattcttaagatgtacaaacccagat
attattaaacaacatgcggctaaagtttacagtaaagagcctcaaggtaagccaccaatg
35  actgtaccacaccttgatacacgttatatcaatggtaaacaaacattattatttggtcca
tatgcgaatatcggccctaaattcttgaaattcggttcaaatctagacttattcgaatca
atcaaaccatataacattactacaatgttggcttcagcagttaaaaatgtacctttaatt
aaatattcaattgatcaaatgatcaaaactaaagaaggttgtatgaactatttaagaaca
tttattcctgatgctaaagatgaagattgggaactttacactgctggtaaacgtgttcaa
40  gttattaaagataaccatttgaagcaggaccttactttattgcagcagtcgtatttgtta
tag


Sequence 3300
MHMSEANHKNIVVVGAGIIGTSVATMLSKVSPNWHIDMFERLEGAGIESSNENNNAGTGH
45  AALCELNYTVEQDDGSIDASKAQEINEQFELSRQFWGNLVKNGDISNPEEFIQPLPHISF
VMGPTNVNFLRKRYETLRTLPMFDTIEYTEDMETMRKWMPLMMENREPGHQMAASKIDEG
TDVNYGALTRKLAHYLEQKSNVSLKYNHDVVDLTQREDGKWEVVVENRETKEKVTKIADK
VFIGAGGHSIPLLQKSGVKQREHLGGFPISGQFLRCTNPDIIKQHAAKVYSKEPQGKPPM
TVPHLDTRYINGKQTLLFGPYANIGPKFLKFGSNLDLFESIKPYNITTMLASAVKNVPLI
50  KYSIDQMIKTKEGCMNYLRTFIPDAKDEDWELYTAGKRVQVIKDNHLKQDLTLLQQSYLL
*


Sequence 3301
Contig_0795_pos_4741_5160
55  is similar to (with p-value 7.0e-17)
>sp:sp|P77279|YBBL_ECOLI HYPOTHETICAL ABC TRANSPORTER ATP-B
INDING PROTEIN IN USHA-TESA INTERGENIC REGION. >gp:gp|U82664
|ECU82664_88 Escherichia coli minutes 9 to 11 genomic sequen
ce. NID: g1773084. >gp:gp|AE000155|AE000155_6 Escherichia co

li K-12 MG1655 section 45 of 400 of the complete genome. NID
: g1786692.
atgcaacaaagtgagttaatcggttatacaattgaagataatatgaaatttcctgctgag
gctagaagtgaagcttttgaccgtgataaagcgaaacaactcatctctcaagtaggatta

5    ggtaattatcagttagatgctcaaattgagcacatgtctgggggagagcaacaacgtatt
accatcgctagacaactcatgtatgaacctgaagtttttattattggacgaagctactagc
gctttagatacacataataaaaagaaaattgaagaaattatatttaaactagcagataaa
gggattgccattttgtggattacgcatagtgatgaccaaagtatgcgtcattttaagcgt
agaatcacaattactgacggtaagatatcgagtgatgaggagttgaatggtaatgagtaa

10

Sequence 3302
MQQSELIGYTIEDNMKFPAEARSEAFDRDKAKQLISQVGLGNYQLDAQIEHMSGGEQQRI
TIARQLMYEPEVLLLLDEATSALDTHNKKKIEEIIFKLADKGIAILWITHSDDQSMRHFKR
RITITDGKISSDEELNGNE*

15

Sequence 3303
Contig_0795_pos_5525_5929
is similar to (with p-value 2.0e-32)
 >sp:sp|P77307|YBBM_ECOLI HYPOTHETICAL 28.2 KD PROTEIN IN US
20   HA-TESA INTERGENIC REGION.
atgttagctaataatggtttaatcgcgattaatctcgcttatcagaatttagaaagagca
tttgttcaagatgtttctgatattgaatccaaacttacgttagcagcgacacctaagctc
gcatcaaaatcagctattagagaaagtatacgcttagcaattgttcctacaattgattct
gtaaaaacatatggtctagttttcaattccaggtatgatgacaggattgattatcggaggc
25   gttgacccacttcaagcaattaaatttcaattgcttgtcgtgtttattcatacaacagcg
acgattatgtctgcactcattgcaacgtatatgagttacggtcagttctttaatgctcgt
catcaactcattgctagaacgcaacgcacaagacaaagtagttaa

Sequence 3304
30   MLANNGLIAINLAYQNLERAFVQDVSDIESKLTLAATPKLASKSAIRESIRLAIVPTIDS
VKTYGLVSIPGMMTGLIIGGVDPLQAIKFQLLVVFIHTTATIMSALIATYMSYGQFFNAR
HQLIARTQRTRQSS*

Sequence 3305
35   Contig_0795_pos_6068_7246
is similar to (with p-value 7.0e-24)
 >gp:gp|A08113|A08113_1 Synthetic DNA sequence of chloramphe
nicol-resistance gene. NID: g413362.
atgactgtcgcacgcacaacaacatttatattaagtgttttcatcgtaggtatggttgag
40   atgatggttgcaggtattatgaatttaatgagtcaagatcttcatgtttctgaagcggtg
gttgggcaactggtaacgttgtacgcgcttacctttgcgatatgtggtccgttattagtg
aagttgactcatcggtttacttcgcgatcagtattattatggacgttaattgtctttatc
ttcgggaatggtatgattgccattgcacctcattttggaataatagttgtaggacgtatt
ttatcttctgccgcagcttcactcattattgtgaaagttttagcactcacagcgatgctc
45   acatcagcaaaaaacagaggtaaaatgattggtattgtttatacaggttttagtggggcg
aatgtctttggtgttcccatcggtacagtgattggcgactgggtaggatggcgatttaca
tttttattcattattattgtaagtgtatttgttggtgttttaatgttaatttatctacca
aaagaagatgaattgtcacatccaaatcaaacacctcgttcatctagtattgaatcacaa
actggctcaagcgtcataagacctcgtgaggttttaaatatctgatgattacatttta
50   gtgctggttgctaattctgtaacattcgtgtttattaatccattaatttttatccaatgga
catgaaatgtctttgtgtgtctttagcactacttgttaatggtgtagcaggtgtgattggt
acttcattaggcggtgtttttatctgataagtttactagtaagcgttggttaataatttcg
atttcaatatttataataatgatgataattcttaacttattattaccaggaacaggatta
ttattagttggcttatttatgtggaatttaatgcagtggagtacaaatccagctattcaa
55   agtggtattattgaacacgtcgaaggagatacaagccaagtgatgagttggaatatgtcg
agtctcaatgccggtattggcgttggcggaatcgtaggtggactcgttatgacacattta
tcagtggaatatgttacttatactagtgcattgataggtttaattagtcttatcattgta
ttcactttaaaaaaatagacattatgctaaaaatttatga

Sequence 3306
MTVARTTTFILSVFIVGMVEMMVAGIMNLMSQDLHVSEAVVGQLVTLYALTFAICGPLLV
KLTHRFTSRSVLLWTLIVFIFGNGMIAIAPHFGIIVVGRILSSAAASLIIVKVLALTAML
TSAKNRGKMIGIVYTGFSGANVFGVPIGTVIGDWVGWRFTFLFIIIVSVFVGVLMLIYLP
KEDELSHPNQTPRSSSIESQTGSSVIRPREVFKYLMITFLVLVANSVTFVFINPLILSNG
HEMSFVSLALLVNGVAGVIGTSLGGVLSDKFTSKRWLIISISIFIIMMIILNLLLPGTGL
LLVGLFMWNLMQWSTNPAIQSGIIEHVEGDTSQVMSWNMSSLNAGIGVGGIVGGLVMTHL
SVEYVTYTSALIGLISLIIVFTLKNRHYAKNL*

Sequence 3307
Contig_0799_pos_3893_3078
is similar to (with p-value 1.0e-20)
 >gp:gp|U67964|EVU67964_2 Ectromelia virus H14-B and H14-E g
enes, complete cds. NID: g2145123.
atgcattatataaaatttattgagtcaaaagataatacaaaactttatatgaaagtgaat
gatattcaagatgcaaaagcgaatatcattatagctcatggtgtggcagaacatttagat
cgttatgatgagataacagcatatttaaatgaagcgggtttttagtgttattagatatgat
caaagagggcatggtcgttctgaaggcaagcgtgccttttatagcaatagtaatgaaatt
gtcgaagatttagatgcgataataaattatgtgaagtcaaactttgaaggtaaagtttac
ttaatcggtcatagtatgggtggttatacagtcactttatatggaacgaaacatccaaat
acagtgaatggtattataacttctggagcattaacacgttataataataaactatttggc
aatcctgatagaaacatatcacctgatacttatatagaaaacaatttaagtgaggggta
tgttctgatttagaggtaatggaaaaatataaacttgatgatttgaatgcgaaacaaatc
tctatggggctcgtcttttcaataatggatggtgttaggtatttgaaagacaatgctcaa
caatttacagataatattttgatattgcatggcaaggaagatgggctagtaagctatgta
gattctttacagctttatcaagaaataggatcagcacataaatcattacacatctatgat
cgtttggagcatgaaatatttaatgaaagttcttataatagaactatttttaacgaagtt
attgaatggcttgaaacggaattaacttataactaa

Sequence 3308
MHYIKFIESKDNTKLYMKVNDIQDAKANIIIAHGVAEHLDRYDEITAYLNEAGFSVIRYD
QRGHGRSEGKRAFYSNSNEIVEDLDAIINYVKSNFEGKVYLIGHSMGGYTVTLYGTKHPN
TVNGIITSGALTRYNNKLFGNPDRNISPDTYIENNLSEGVCSDLEVMEKYKLDDLNAKQI
SMGLVFSIMDGVRYLKDNAQQFTDNILILHGKEDGLVSYVDSLQLYQEIGSAHKSLHIYD
RLEHEIFNESSYNRTIFNEVIEWLETELTYN*

Sequence 3309
Contig_0802_pos_4926_3982
is similar to (with p-value 3.0e-73)
 >gp:gp|S72926|S72926_1 Hordeum vulgare glucose and ribitol
dehydrogenase homolog mRNA, complete cds. NID: g633889.
gtgtgggcttggtacactcatttacaagaattcaattctgatccgaatatacaaaattt
gatgaaatgcttaacaaactacaaaaggtcagtttaattagtgcaagtgaagaaagtggg
actaaaaaaattgtagatcactttgtcgaagaattatatagcgaagaaccaaaacaaaaa
atcaatacaggttataaactggtggattacaaaataggtggtttagaacctacacagttg
attgtaatcgctgcgagaccgtcagtaggtaaaacggggtttgcgcttaatatgatgctt
aatatagcgtctcaaggctataaaacttcattcttcagtctagagacaactggcgtgtct
gtattgaaaaggatgttatcagcagaaactgggatagaactaactcgtatcaaagaaatt
aaagatttagaaccggatgaattaacacgtttaacaactgcagcagacagaatactcaaa
cttgatatagatatacacgataaaagcaatattactacacatgatgtacgtaaacaagcg
atgaagaacaaagatgtgcaacaggttatcttcattgactacttacaacttatgcagaca
gacagtaagttagatcgtcgtaatggtatcgaaaagatatcgcgagatttgaagattatt
gcaaatgaaacaggtgcaattattgtgttgctatctcaattgagcagaggtgtagaaaca
agaaatgacaaaagacctatgctatctgacatgaaagaagcaggtggaattgaagcagat
gcaagtttagctatgttgttatatcgagatgattactacaaccgtgatgatgttgatgac
tcaggcaagtcaattgttgaatgtaacatcgcaaagaataaagacggagaaacaggtgta
gttgagtttgagtactacaagaaaacgcagaggttcttcacatga

Sequence 3310

VWAWYTHLQEFNSDPNIQNFDEMLNKLQKVSLISASEESGTKKIVDHFVEELYSEEPKQK
INTGYKLVDYKIGGLEPTQLIVIAARPSVGKTGFALNMMLNIASQGYKTSFFSLETTGVS
VLKRMLSAETGIELTRIKEIKDLEPDELTRLTTAADRILKLDIDIHDKSNITTHDVRKQA
MKNKDVQQVIFIDYLQLMQTDSKLDRRNGIEKISRDLKIIANETGAIIVLLSQLSRGVET
RNDKRPMLSDMKEAGGIEADASLAMLLYRDDYYNRDDVDDSGKSIVECNIAKNKDGETGV
VEFEYYKKTQRFFT*

Sequence 3311
Contig_0802_pos_3039_2170
is similar to (with p-value 3.0e-36)
>sp:sp|P37469|DNAC_BACSU REPLICATIVE DNA HELICASE (EC 3.6.1
.-). >gp:gp|D26185|BAC180K_4 B. subtilis DNA, 180 kilobase r
egion of replication origin. NID: g467326. >gp:gp|Z99124|BSU
B0021_149 Bacillus subtilis complete genome (section 21 of 2
1): from 3999281 to 4214814. NID: g2636442.
gtgaatttattaaaatttcacaacaaaatcaaaggatatactcaaaatagacaaccaggt
attgaagcggatatggaacctaaacccattgcagaattagaagaatataaagcagcagga
aagttagagaataaagttgctctaataacaggaggagattcaggtattggacgtgcgata
gcaatactatatgctaaagaaggggcaaatgttgctattggttattatgacgaacatcaa
gatgccgaagacacagttaatcgacttcaagaaatgggtgtaaaagctaaagcttatgct
catgacttaaaagatgaaaagcaatctcaaaagttaatcaaagatgtcataaatgacttc
ggtagtttaaatatattagtaaataatgggggcgtgcaatttccacgcgatcattttgaa
gatatcactccacaacaagtgaaagaaacttttatgacgaatattttttggtatgatgttt
ttatcccaatcagcagtaccttacctatctgaaggagatacaattataaacactacaagt
gtcacagcatatagaggatcggggcatctcattgattattcagctacaaaaggtgccata
gtatcgtttacccgttctcttgctactactttaatggaaaagggaattcgcgttaacgcc
gttgcgcccggcccaatctattcacctttaattcctgcgacatttgatgaagaaaaagta
gaacatcaaggtgacgaaacgccgatgggtcgtcgtggacaaccagcagaacttgcacct
tcttatgtcttcttagcaacacatgcagatagttcctatattactggtcaagtcattcat
gtcaatggtggcgattttatcacatcttaa

Sequence 3312
VNLLKFHNKIKGYTQNRQPGIEADMEPKPIAELEEYKAAGKLENKVALITGGDSGIGRAI
AILYAKEGANVAIGYYDEHQDAEDTVNRLQEMGVKAKAYAHDLKDEKQSQKLIKDVINDF
GSLNILVNNGGVQFPRDHFEDITPQQVKETFMTNIFGMMFLSQSAVPYLSEGDTIINTTS
VTAYRGSGHLIDYSATKGAIVSFTRSLATTLMEKGIRVNAVAPGPIYSPLIPATFDEEKV
EHQGDETPMGRRGQPAELAPSYVFLATHADSSYITGQVIHVNGGDFITS*

Sequence 3313
Contig_0804_pos_4400_3708
is similar to (with p-value 3.0e-62)
>sp:sp|P96612|DDL_BACSU D-ALANINE--D-ALANINE LIGASE (EC 6.3
.2.4) (D-ALANYLALANINE SYNTHETASE). >gp:gp|AB001488|AB001488
_40 Bacillus subtilis genome sequence, 148 kb sequence of th
e region between 35 and 47 degree. NID: g1881226. >gp:gp|Z99
106|BSUB0003_103 Bacillus subtilis complete genome (section
3 of 21): from 402751 to 611850. NID: g2632653.
atggataaactcgtgatgaaacaattatttgagcatagaggtttacctcaattaccttat
attagctttttaagaagtgagtatgaaaaatatgaaaataatatcattaaaattagttaat
gataagttaacatatccggtatttgtaaaacctgctaatctcggttcaagtgttggtata
agtaaatgtaacaatgaagaagaattaaaatctgggatagctgaagcattccaatttgat
cgtaaacttgtcattgaacaagggattaatgctagagagatcgaagtagctgtcttaggt
aacgattatcctgaaacgacatggcctggtgaggttgttaaggatgtagcgtttttatgat
tataaatcaaagtataaagacggtaagattagattagatattccagcagatttagatcaa
gatgttcaaatgacattaagaaacatggcattagaggcctttaaagctactgattgttcg
ggattagttcgtgcagatttctttgttactgatgataatcaaatttatattaatgaaaca
aatgccatgccaggatttactgcatatagtatgtatccaaacttatggaaaaaatatgggc
ttatcttaccctgatttaattgctaaattgattgatttagctaaagaacgttatgaagat
aaaaagaaaaataaatataaaattgattattag

Sequence 3314
MDKLVMKQLFEHRGLPQLPYISFLRSEYEKYENNIIKLVNDKLTYPVFVKPANLGSSVGI
SKCNNEEELKSGIAEAFQFDRKLVIEQGINAREIEVAVLGNDYPETTWPGEVVKDVAFYD
5    YKSKYKDGKIRLDIPADLDQDVQMTLRNMALEAFKATDCSGLVRADFFVTDDNQIYINET
NAMPGFTAYSMYPNLWKNMGLSYPDLIAKLIDLAKERYEDKKKNKYKIDY*


Sequence 3315
Contig_0808_pos_6745_6269
10   is similar to (with p-value 1.0e-55)
        >gp:gp|AF068246|AF068246_1 Mus musculus SA protein mRNA, co
     mplete cds. NID: g3928675.
     gtgtttaattacttatctacgaaagaagacgaaagagaatgggttgaagcaattagagta
     gcaagaaatatcctaaaacaaaaagctatggacccatttaatggtggcgaaatttcacca
15   ggaccacaagttcaaacggatgaagaaattctagattgggtacgtaaagatggagaaact
     gcattacatccatcttgtagcgcgaaaatgggacctgcatctgacccaatggcagtagtc
     gatccattaactatgaaagtacatggtatggaaaatttacgtgtcgttgatgcttcagca
     atgcctagaacaacaaatggtaatattcatgcacctgtattgatgttagctgagaaagca
     gcggacattattcgtggtagaaaaccgcttgaacctcaatatgttgactattataaacat
20   ggtattgatgatgaaaaagcaggtgcaatggaagatgatccattctaccaatattaa


Sequence 3316
VFNYLSTKEDEREWVEAIRVARNILKQKAMDPFNGGEISPGPQVQTDEEILDWVRKDGET
ALHPSCSAKMGPASDPMAVVDPLTMKVHGMENLRVVDASAMPRTTNGNIHAPVLMLAEKA
25   ADIIRGRKPLEPQYVDYYKHGIDDEKAGAMEDDPFYQY*


Sequence 3317
Contig_0808_pos_3296_2034
is similar to (with p-value 7.0e-29)
30      >sp:sp|P17444|BETA_ECOLI CHOLINE DEHYDROGENASE (EC 1.1.99.1
     ) (CHD). >gp:gp|X52905|ECBET_5 Escherichia coli betT, betI,
     betB and betA genes. NID: g48714. >gp:gp|M77738|ECOBETA_1 E.
     coli choline dehydrogenase (betA) gene, complete cds. NID: g
     145401. >gp:gp|AE000138|AE000138_2 Escherichia coli K-12 MG1
35   655 section 28 of 400 of the complete genome. NID: g1786501.

     atgaataaatcaaatttactagcacctgagaattataatattgttacagaaatagaaaaa
     tatgcctcagaagatcataaaaaagccattatttacaaggataacgagcatgaaaatatt
     tctgtaagttataaagaacttatcagtaatgctaataaagtagggaatgtattcctcaat
40   catgggctaaaaaagggagataaagttctcatcatgatgccacgtgcaatcgttacatat
     gaattatatattgcagcattgaaactagggatagcgattgttccaagttcggaaatgtta
     cgaacaaaagatttacaatatcgaattactcacggtgagattgatgcagttatttcattt
     gattctctaactaaagaatttgaaaacgttaaagaatatgaccaattaaaaaaatttata
     gtagctggtcacaaagaagattgggtttcaatagaagatgaaaaagaaaaagtaagtgat
45   gaccttaaaggcgcagatacaacacgagatgatttggcgattctttcttatacatcaggt
     acaacaggcaatccaaaagcagtaacgcattcacatggatgggggtatgcccatttacaa
     atggcaccaaaacattggttatgtatacaagagaatgatcttgtatgggcaactgcagca
     ccagggtggcaaaagtgggtgtggagtccatttttatctgtattagggatgggagcaaca
     gcatttgtctataacggtcgtttccaccctgaaacatatctcgagttacttcaaaattat
50   caaattaatgttctatgttgtacaccaacagaatatcgtatgatggctaaacttagtcat
     ttagaacagtacaatttagagtatttacacagtgcggtgtctgcgggtgaacctttaaat
     cgagaagttgttgaacaatttaaacgtcattttaatattactgttcgagatggatatgga
     caaaccgaaagtacattgttgatcggatttctaaaagatactgaaccacgtatgggttct
     atgggcaaaggtatacctggtagttttgttactgtcattgacgatgatggtaaagaggtt
55   ggtccaaatgttaaaggtaatatcgccgtgcctttagacttaccggctttatttaaaggt
     tactttaaagatgaagcacgcacaaaagcagcttcaacaggtgattattatgttactgga
     gaccaagctcatattgatatatatcagtggtcaatcatttttacgtcctttattttaaaa
     taa


830

Sequence 3318
MNKSNLLAPENYNIVTEIEKYASEDHKKAIIYKDNEHENISVSYKELISNANKVGNVFLN
HGLKKGDKVLIMMPRAIVTYELYIAALKLGIAIVPSSEMLRTKDLQYRITHGEIDAVISF
DSLTKEFENVKEYDQLKKFIVAGHKEDWVSIEDEKEKVSDDLKGADTTRDDLAILSYTSG
5      TTGNPKAVTHSHGWGYAHLQMAPKHWLCIQENDLVWATAAPGWQKWVWSPFLSVLGMGAT
AFVYNGRFHPETYLELLQNYQINVLCCTPTEYRMMAKLSHLEQYNLEYLHSAVSAGEPLN
REVVEQFKRHFNITVRDGYGQTESTLLIGFLKDTEPRMGSMGKGIPGSFVTVIDDDGKEV
GPNVKGNIAVPLDLPALFKGYFKDEARTKAASTGDYYVTGDQAHIDIYQWSIIFTSFILK
*

10

Sequence 3319
Contig_0810_pos_827_1459
is similar to (with p-value 8.0e-31)
 >gp:gp|D64024|D64024_2 Sulfolobus sp. DNA for 2-oxoacid:fer
15     redoxin oxidoreductase subunit alpha and beta, complete cds.
NID: g1565182.
atggcaaacaaagatttaacagttatcgcttctggtggtgatggagacggctatgcaata
ggaatgggacatactattcatgctcttagacgtaatatgaatatgacgtatattgtcatg
gacaatcaaatatatggattaactaaaggacaaacatcaccttcctcagctaaaggattt
20     gtaactaaatcaacacctaaaggaaatatagaaaagaatgtagctccattggaattggca
ctgtcctctggtgcaacttttgtagcacaaggattctcaagtgatataaaggcattaact
aaaatgattgaagatgcgattcatcatgatggttttttctttttgttaatgttttctcacct
tgtgttacttacaataaagtgaatacttatgactggtttaaagaacatttaacaagtatc
gatgatattgagggctatgacatcacagataaacaacttgctatgaaaactgtgctggat
25     catgagtcactggttaaaggtatcgtttatcaagatacaacaacaccttcttatgaatcg
caaatttcagaactagaacatgaggcgttagctaaaagagatattcatattacagaagaa
actttcaacgatttaactgcacaattttttataa

Sequence 3320
30     MANKDLTVIASGGDGDGYAIGMGHTIHALRRNMNMTYIVMDNQIYGLTKGQTSPSSAKGF
VTKSTPKGNIEKNVAPLELALSSGATFVAQGFSSDIKALTKMIEDAIHHDGFSFVNVFSP
CVTYNKVNTYDWFKEHLTSIDDIEGYDITDKQLAMKTVLDHESLVKGIVYQDTTTPSYES
QISELEHEALAKRDIHITEETFNDLTAQFL*

35     Sequence 3321
Contig_0812_pos_620_1003
is similar to (with p-value 3.0e-63)
 >pir:pir|A43577|A43577 regulatory protein pfoR – Clostridiu
m perfringens
40     atgactgctaatgcatatgggttgttatcacctactgaagtgattgcactaccaatttgt
ttaagtactgcaataacccctggagaaattaagtttgcaagaccaaatgtaactgctgga
gccaataaaataacaacaattaaatctagtccctctggaactttttctcaacaaacttg
attccaaatgcaacaacatacgctgcgataaacgcaggtaataattttgagtcatgtaat
actaaacctacaatgactgcaaatactggagacacttctaactttagacaagttaatata
45     ccaacagctataccacttaaactccctgctagatcaccaatttcttggaaaaatttaaca
tggaatacgccaccaattgcgtaa

Sequence 3322
MTANAYGLLSPTEVIALPICLSTAITPGEIKFARPNVTAGANKITTIKSSPSGTFFSTNL
50     IPNATTYAAINAGNNFESCNTKPTMTANTGDTSNFRQVNIPTAIPLKLPARSPISWKNLT
WNTPPIA*

Sequence 3323
Contig_0812_pos_2366_5200
55     is similar to (with p-value 2.0e-07)
 >sp:sp|O34863|UVRA_BACSU EXCINUCLEASE ABC SUBUNIT A. >gp:gp
|Z99122|BSUB0019_13 Bacillus subtilis complete genome (secti
on 19 of 21): from 3597091 to 3809700. NID: g2636029. >gp:gp
|AF017113|AF017113_13 Bacillus subtilis 300-304 degree genom

ic sequence. NID: g2618830.
atgaaaggaccgtcaattgtagtaaaaggtgcacgagcacataatttaaaaggagtagat
attgaattacctaaaaacaagttgattgtcatgactggactttcaggctcaggtaaatct
tcgttagcgttcgatactatttatgctgaaggtcaacgacgttatgtggaatcgttaagt
gcatatgcgcgacaattttttaggacaaatggacaaacctgatgtagataccattgaaggt
ttgtcaccagcaatatctattgatcagaagacgacaagtaaaaacccacgttctaccgtt
gccactgtgactgaaatctatgattatattcgattattatatgcgcgtgtcggtaaaccg
tactgtccttatcacggtatagaaattgaatcacaaactgttcaacaaatggttgatcgt
attttagaattggaagaacgtacaaaaattcaattacttgctccagttatatcacataga
aaaggaagtcatgaaaaattaattgaagatataggtaagaaaggatatgtgcgtttacgt
gttgatgatgaaatcgtcgatgtaaatgaagtgcctcaattggataaaaataaaaatcac
acaattgaagttgttgttgatagacttgtagttaaagatggcattgagacgcgtttagcg
gactctattgagacagcacttgaacttgctgaaggtaatttaacagttgacgttattaat
ggagaagaacttaaattctctgagaaccacgcatgtcctatatgtggtttctctattggt
gaattagaacctagaatgtttagtttcaacagtccgtttggtgcgtgtccaacttgtgat
ggtttaggtcaaaaactaaaagtagacctagatttagtcattcctgataaaaataaaact
ttaaatgaaggtgcaattgaaccatgggaaccaacaagttcggattttttatccaactta
ttaaaacgtgtatgtgaagtctacaaaattaatatggataagccatataaaaaattgact
gatagacaaaaaaatatacttatgaatggttctggagaaaaagaaattgaatttactttc
acgcaacgaaatggtggtactcgtaaacgtaaaatggttttgaaggtgtggtacctaac
attgaccgtcgttaccatgaatctccctctgaatatacacgcgaaatgatgagtaaatat
atgacagagttaccgtgtgaaacttgtcatggcaaacgcttaagcaaagaagcgttatct
gtatacgtaggcgactataatataggtgaagtcgttgaatattcaattaaaaacgcgctt
tattatttcgaaaatttaaaattaagcgaccaagatataatcgattgcagatcaaattta
aaagagattatttcaagattatcattttaaataatgttggtttggaatatttgacttta
gatcgctcatcagggactttatctggaggggaagcacaacgtattcgtttagcaacacaa
attggttcacgattaaccggagtactatatgtttagatgaaccttctattggattacat
caaagagataatgatagattaattaatactttaaaagaaatgcgtgatttaggtaataca
cttattgtcgttgaacatgacgatgatactatgagagcagctgattattagttgatgtg
ggtccgggagctggtaaccacggtggagaggttgtctcaagtggtacccctaataaagta
atgaaagataaaaatccttaactggtcaatattaagtggaaaaaacgaattgaagtc
cctgaatacagacgagaaatcaccgatagaaagattcaaattaaaggtgctaaaagtaat
aatttgaaaaatgtaaatgtagacttcccactatctgtcttaactgttgttacaggtgtg
tcaggttctggtaaaagttcactcgtcaatgaaattttatataaagcattagctcaaaaa
attaataaatctaaagtgaaacctgggaattttgatgaaattaaaggaattgatcaatta
gataaaatcattgatattgatcaatcgccaataggtagaacaccacgttctaacccagcc
acatacactggtgtctttgatgacataagagatgtctttgcacaaacgaatgaagctaaa
atacgaggttatcaaaaaggtagatttagtttaatgtcaaaggtggacgatgtgaagct
tgtaaaggtgatggaattataaaaattgaaatgcatttttaccagatgtctatgtacct
tgtgaagtatgtgatggtaaacgctataatcgtgagactttagaggtaacatacaaaggt
aaaaatattgcggatgtattagaaatgactgttgaagaagctacgcatttctttgaaaat
attcctaagattaaacgtaaattacaaacacttgtagatgttgggttggggtacattact
ttaggtcaacaagctactacattatctggtggcgaagcgcaacgtgtaaaactcgcatca
gaattgcacaaacgttcaacggggcgttctatttatattcttgatgaaccaactacagga
ttacatgtcgacgatataagtcgtttattaaaggtattgaatcgtatagtggaaaatggt
gatacggtcgttattatcgaacacaatcttgatgttattaaaacggctgatcatattatt
gatttaggtccagaaggcggtgaaggtggaggaacaatcatcgcaactggtacacctgaa
gagattgctcaaaataaagggtcttacactggtcaatacttaaaaccagtattagagaga
gacagcgttgaatag

Sequence 3324
MKGPSIVVKGARAHNLKGVDIELPKNKLIVMTGLSGSGKSSLAFDTIYAEGQRRYVESLS
AYARQFLGQMDKPDVDTIEGLSPAISIDQKTTSKNPRSTVATVTEIYDYIRLLYARVGKP
YCPYHGIEIESQTVQQMVDRILELEERTKIQLLAPVISHRKGSHEKLIEDIGKKGYVRLR
VDDEIVDVNEVPQLDKNKNHTIEVVVDRLVVKDGIETRLADSIETALELAEGNLTVDVIN
GEELKFSENHACPICGFSIGELEPRMFSFNSPFGACPTCDGLGQKLKVDLDLVIPDKNKT
LNEGAIEPWEPTSSDFYPTLLKRVCEVYKINMDKPYKKLTDRQKNILMNGSGEKEIEFTF
TQRNGGTRKRKMVFEGVVPNIDRRYHESPSEYTREMMSKYMTELPCETCHGKRLSKEALS
VYVGDYNIGEVVEYSIKNALYYFENLKLSDQDKSIADQILKEIISRLSFLNNVGLEYLTL

DRSSGTLSGGEAQRIRLATQIGSRLTGVLYVLDEPSIGLHQRDNDRLINTLKEMRDLGNT
LIVVEHDDDTMRAADYLVDVGPGAGNHGGEVVSSGTPNKVMKDKKSLTGQYLSGKKRIEV
PEYRREITDRKIQIKGAKSNNLKNVNVDFPLSVLTVVTGVSGSGKSSLVNEILYKALAQK
INKSKVKPGNFDEIKGIDQLDKIIDIDQSPIGRTPRSNPATYTGVFDDIRDVFAQTNEAK
5    IRGYQKGRFSFNVKGGRCEACKGDGIIKIEMHFLPDVYVPCEVCDGKRYNRETLEVTYKG
KNIADVLEMTVEEATHFFENIPKIKRKLQTLVDVGLGYITLGQQATTLSGGEAQRVKLAS
ELHKRSTGRSIYILDEPTTGLHVDDISRLLKVLNRIVENGDTVVIIEHNLDVIKTADHII
DLGPEGGEGGGTIIATGTPEEIAQNKGSYTGQYLKPVLERDSVE*

10    Sequence 3325
Contig_0812_pos_1152_109
is similar to (with p-value 2.0e-22)
    >pir:pir|A43577|A43577 regulatory protein pfoR - Clostridiu
m perfringens
15    atggatattattttaggagtagggactttagtactcgttcttattatcatgacgctttc
ttaaattttgcgccatatggtaaacaaggtttacaagctttatcaggggctgcttgtgcc
acgtttttaccacaggcgttcttaagttacgcaattggtggcgtattccatgttaaattt
ttccaagaaattggtgatctagcaggggagtttaagtggtatagctgttggtatattaact
tgtctaaagttagaagtgtctccagtatttgcagtcattgtaggtttagtattacatgac
20    tcaaaattattacctgcgtttatcgcagcgtatgttgttgcatttggaatcaagtttgtt
gagaaaaaagttccagagggactagatttaattgttgttattttattggctccagcagtt
acatttggtcttgcaaacttaatttctccaggggttattgcagtacttaaacaaattggt
agtgcaatcacttcagtaggtgataacaacccatatgcattagcagtcattttaggactt
gttattcctgtaactggtatgacgccattaagctcaatggtacttacaagcttattaggt
25    ttaactggtattccaatggcaattggtgcattaacatgtacaggagcatcttttgttaat
ggaatcttatttagcaaattaaaaattggtaataaaggtaatgccttcgcggtatttgta
gaaccgttaactcaaattgacttaattgccaaatatccactacaactgtttggtgcgaat
gccattattggtgttgtaaatgcttgtattgtcacatacagtggactaattattgatatt
aaaggtatggcaacacctatagcaggtgctattgtactttatggctttaacgacgctgta
30    agatctacaattacaattatcgcagtagcaattgcaagtgtgatattagcgtacgttatt
agtgctattattaataaatttaacttgatgaatgtcggattcaagttaccacgtagaaaa
aaccaagttaaggagagtgtttaa

Sequence 3326
35    MDIILGVGTLVLVLIIMTLFLNFAPYGKQGLQALSGAACATFLPQAFLSYAIGGVFHVKF
FQEIGDLAGSLSGIAVGILTCLKLEVSPVFAVIVGLVLHDSKLLPAFIAAYVVAFGIKFV
EKKVPEGLDLIVVILLAPAVTFGLANLISPGVIAVLKQIGSAITSVGDNNPYALAVILGL
VIPVTGMTPLSSMVLTSLLGLTGIPMAIGALTCTGASFVNGILFSKLKIGNKGNAFAVFV
EPLTQIDLIAKYPLQLFGANAIIGVVNACIVTYSGLIIDIKGMATPIAGAIVLYGFNDAV
40    RSTITIIAVAIASVILAYVISAIINKFNLMNVGFKLPRRKNQVKESV*

Sequence 3327
Contig_0814_pos_2358_4124
    >sp:sp|O67589|SYD_AQUAE ASPARTYL-TRNA SYNTHETASE (EC 6.1.1.
45    12) (ASPARTATE--TRNA LIGASE) (ASPRS). >gp:gp|AE000750|AE0007
50_9 Aquifex aeolicus section 82 of 109 of the complete geno
me. NID: g2983999.
atgaataaaagaacaacgtattgtggtttagtcacagaagaattttttaaacgaaaaagta
acattaaaaggttgggttcataacagacgagatttaggtggattaattttttgttgatttta
50    agagatcgtggaggtattgtccaaatagtttttaatcctgacttttccgaagaagcattg
caagttgctgaaacagtacgctcagaatatgtagtcgaagttgaaggtgtagtaacaaaa
cgtgatgctgaaactattaacccaaaaatcaaaacaggtcaagttgaggttcaagtttca
aatattgagattattaacaaatcagaaacacctccattttcaattaatgaagaaaatgta
aacgttgatgaaaatattcgattaaaatatagatatttagatttacgtagacaagaatta
55    gcgcaaacttttaaaatgagacatcaaactacgcgttctatccgtcaatacttagataat
aatggcttcttcgatattgaaacaccagtattaacaaaatcaacacctgagggtgcgcga
gattatctagtaccttcccgtgtacacgagggtgaattttacgcgttgccacaatcacca
caattatttaaacaactattaatgataagtggttttgataaatattatcaaattgttaaa
tgtttccgtgatgaagacttacgtgcagatcgtcaaccagaattcactcaagttgatatt

```
gaaatgagttttgtcgatcaagaagatatcatagccatgggtgaagatatgttacgtaag
gttgtaaaagatgtaaaaggaatagacgttagtggcccattcccacgtatgacatatgca
gaggctatggaccgttttggttcagataaacctgacactcgtttcggtatggaacttatc
aatgtgtcacagcttggtaaagaaatgaattttaaagtttttaaagatacggtagataac
aacggcgaaattaaagcaattgtcgcaaaagacgctgcaaataaatatacacgtaaagac
atggatgcattaacagagtttgtaaatatatatggtgcaaaaggattagcttggggttaaa
gttgttgatgatggttttaagtggcccaattgctagattttttcgaagatgttaatgttgaa
acacttaaacagttaacagaagctaaacctggagatttagtaatgtttgtagctgataaa
cctaatgttgttgctcaaagtttaggggctttaagaattaaattagcaaaagaattaggt
ttaattgatgaatcaaaattaaattcttatgggtaactgattggccgttattagagtat
gatgaagatgcaaaacgttatgtagcagcacatcatccatttacttcacctaaaagagaa
gatatcgaaaagctagacactgaacctgaaaatgtacaagccaacgcttatgatattgtt
ctaaatggttatgaacttggtggtggttctataagaatacacgatggtgaattgcaacaa
aaaatgtttgaagtattaggatttactaatgaacaagtcaagaacaattggtttctta
ttagatgctttttaaatacggtgctccacctcatggtggcatcgcgttaggtttagataga
cttgtgatgttattaacaaatagaacaaacttgagagatacaattgcattccctaaaaca
gcatcagctacatgtctttttaactgacgctccaggagaagtatctgataaacaactccaa
gaactctcactaagaatcagacactag
```

Sequence 3328
MNKRTTYCGLVTEEFLNEKVTLKGWVHNRRDLGGLIFVDLRDRGGIVQIVFNPDFSEEAL
QVAETVRSEYVVEVEGVVTKRDAETINPKIKTGQVEVQVSNIEIINKSETPPFSINEENV
NVDENIRLKYRYLDLRRQELAQTFKMRHQTTRSIRQYLDNNGFFDIETPVLTKSTPEGAR
DYLVPSRVHEGEPYALPQSPQLFKQLLMISGFDKYYQIVKCFRDEDLRADRQPEFTQVDI
EMSFVDQEDIIAMGEDMLRKVVKDVKGIDVSGPFPRMTYAEAMDRFGSDKPDTRFGMELI
NVSQLGKEMNFKVFKDTVDNNGEIKAIVAKDAANKYTRKDMDALTEFVNIYGAKGLAWVK
VVDDGLSGPIARFFEDVNVETLKQLTEAKPGDLVMFVADKPNVVAQSLGALRIKLAKELG
LIDESKLNFLWVTDWPLLEYDEDAKRYVAAHHPFTSPKREDIEKLDTEPENVQANAYDIV
LNGYELGGGSIRIHDGELQQKMFEVLGFTNEQAQEQFGFLLDAFKYGAPPHGGIALGLDR
LVMLLTNRTNLRDTIAFPKTASATCLLTDAPGEVSDKQLQELSLRIRH*

Sequence 3329
Contig_0814_pos_4680_5456
is similar to (with p-value 6.0e-34)
  >sp:sp|Q57097|YGDL_HAEIN HYPOTHETICAL PROTEIN HI0118. >pir:
  pir|C64049|C64049 molybdopterin biosynthesis protein (chlN)
  homolog - Haemophilus influenzae (strain Rd KW20) >gp:gp|U32
  698|U32698_1 Haemophilus influenzae Rd section 13 of 163 of
  the complete genome. NID: g3212178.
```
atgaaacatcaattttcaaggaatgaattagcaataggacaagaagggctgaacttacta
aaaaataagactgttgcagtttaggtgttggtggcgtcgggtcatttgcagctgaggca
ttggctcggactaatatagggcacatcatacttatagataaagatgatgtcgatattaca
aatgtgaacaggcaaattcatgcactgacttcaactattggtcaaagtaaagtcacgcta
atggaagaaagaatcaaattaataaatcccgattgtaaagtaacttcttttgcatatgttt
tataccgaggaaacatacaaagatatcttcaataattatgatattgattattttattgat
gcaagcgatacaatcatttataaagttcatctcatgaaagagtgtttagaaagaggaatt
gagttaatttcaagtatgggtgcagcaaataagactgacccgacacgttttgaaattgca
gatatttcaaaaacacatactgatcctatggctaaagtaattagaaatcgtttaaaacgc
cttggtattcgtaaaggtgttaaagtagtattttctgatgaaagtcctattgttattcgc
gaggacgtaaaagaaacagtaggagataaaaatgcaatcaatagaaaagggcaaatgcct
ccatcttctaatgcatttgttccaagtgtagtaggccttatttgtgcaagctacgttgtc
aacgatattttaaaagatataccgtaaggcgaattaaagataaaggacaaaattaa
```

Sequence 3330
MKHQFSRNELAIGQEGLNLLKNKTVAVLGVGGVGSFAAEALARTNIGHIILIDKDDVDIT
NVNRQIHALTSTIGQSKVTLMEERIKLINPDCKVTSLHMFYTEETYKDIFNNYDIDYFID
ASDTIIYKVHLMKECLERGIELISSMGAANKTDPTRFEIADISKTHTDPMAKVIRNRLKR
LGIRKGVKVVFSDESPIVIREDVKETVGDKNAINRKGQMPPSSNAFVPSVVGLICASYVV
NDILKDIPVRRIKDKGQN*

Sequence 3331
Contig_0815_pos_2100_949
is similar to (with p-value 2.0e-49)

5   >gp:gp|U59234|SPU59234_4 Synechococcus PCC7942 biotin carbo
xylase (accC) gene, complete cds. NID: g2661137.
atgcgtgcagaaaatgaattagtagaacaaaaacaacaagaaaaacaaaaagcattgtat
aaacaagagttagcatggatgcgagcaggagcaaaggcaagaactactaaacaacaggca
cgtatcaatagatttaatcaactagaatcagacgttaagacgcaacatacacaagataag
10 ggtgaacttaatcttgcatattcaaggttaggtaaacaagtatatgaattaaagaattta
tcaaaatcaattaataataaagtttttatttgaagatgtcactgaaattattcaaagtggt
agacgtataggtattgtaggacctaatggagcgggaaaaacaacattacttaatatttta
agtaatgaagatcaggactatgagggtgagcttaaaatcggtcagactgttaaggtagct
tattttaagcaaacagaaaagacacttgaccgtgatattagagtgattgactacctaaga
15 gaagaaagtgaaatggctaaagaaaaagatggtacctcaatttcagttacacaattgtta
gaaagattttttatttccgagcgctacacacggtaaaaaagtttataaactctcaggtgga
gaacaaaaacgtctgtatttattgcgtttacttgttcataaacctaatgtactccttttta
gatgaaccgactaatgatttagatactgaaacacttacgattttagaagattacattgat
gatttcggtggttctgtcattacggtcagtcatgatcgttatttcttaaataaagtggta
20 caagaatattggtttattcatgatggtaaaatcgaaaaaattattggatcatttgaagat
tatgaatcttttaaaaaggaacatgaacgccaagccatgctatctaaacaaactgaacaa
caaaataaacataagcatcaaccaaaaaagaaaacaggactatcttataaagagaagtta
gaatacgaaacaattatgacgcgtatagaaatgactgaaacgcgtttagaagaccttgaa
caagaaatgattaatgcaagtgataattatgcaagaatcaaagaacttaatgaggaaaaa
25 gagcaacttgaagcaacctatgaagcagacatcacgagatggagtgagcttgaggaaatt
aaagaacaataa

Sequence 3332
MRAENELVEQKQQEKQKALYKQELAWMRAGAKARTTKQQARINRFNQLESDVKTQHTQDK
30 GELNLAYSRLGKQVYELKNLSKSINNKVLFEDVTEIIQSGRRIGIVGPNGAGKTTLLNIL
SNEDQDYEGELKIGQTVKVAYFKQTEKTLDRDIRVIDYLREESEMAKEKDGTSISVTQLL
ERFLFPSATHGKKVYKLSGGEQKRLYLLRLLVHKPNVLLLDEPTNDLDTETLTILEDYID
DFGGSVITVSHDRYFLNKVVQEYWFIHDGKIEKIIGSFEDYESFKKEHERQAMLSKQTEQ
QNKHKHQPKKKTGLSYKEKLEYETIMTRIEMTETRLEDLEQEMINASDNYARIKELNEEK
35 EQLEATYEADITRWSELEEIKEQ*

Sequence 3333
Contig_0815_pos_935_528
is similar to (with p-value 1.0e-32)
40   >pir:pir|S30712|S30712 DNA helicase - Escherichia coli
atgcaagagactttatcgcattattttggttataagtcatttcgaccaggacaagaagaa
attataactaaaatattaaatcatcagcatacactgggggtactacctaccggtggcggt
aaatcgatatgctatcaagtaccaggtttaatgcagggtggcacaactattgttataagt
ccacttatttctttgatgaaagaccaagttgatcaactacaagcaatgggaattcaagct
45 gcatatttgaatagtagtttgactcataaacaacaaaaagagattgaagagcaaataaag
cgtggtgccattcagttttatatgtagctccagagcgctttgaaaatactttttttcta
aatttattacgtaaaatagaaatacccccttatcgcttttgagcaataa

Sequence 3334
50 MQETLSHYFGYKSFRPGQEEIITKILNHQHTLGVLPTGGGKSICYQVPGLMQGGTTIVIS
PLISLMKDQVDQLQAMGIQAAYLNSSLTHKQQKEIEEQIKRGAIQFLYVAPERFENTFFL
NLLRKIEIPLIAFEQ*

PU3480
55
1
                               SEQUENCE LISTING

Sequence 3335

```
step.1000b10.cons.ok
TCATACCCATAATGGCAGCACCTTGCATTAATGCGTTAGCAACATTGTTAC
GTCCATCGCCAACATAAGTAAAGTTGATATCAGCATACTCTTTTTTCAATA
CTTCTTTAGCAGTTAAAAAGTCAGCAAGCACTTGTGTAGGGTGATCTTCAT
CGGTTAATCCATTCCATACCGGAACACCTGAATATTGCGCTAATGTTTCTA
CAGTACGTTGAGAGAAACCTCGGTACTCAATACCATCATACATACCACCAA
GTACACGTGCTGTATCTTTAGCAGTTTCTTTTTTACCCATTTGAGAACCTG
TAGGTCCAAGATAAGTGACGTGTGCACCTTGATCATGTGCGGCAACTTCGA
ATGCGCAACGTGTTCTAGTGGAATCTTTTTCAAAAAGAAGCGCGATATTTT
TATTTTTTAGCATAGGCTTTTCAGTGCCGATATATTGGCACGCTTCAAAT
CTTCGGAGAGTGTTAATAAAAATTCTACCTCTTGTCGTGAAAAGTCTAATA
AAGTTAAAAAGCTTCTGTTACGTAAATTTTTCATCAATAACAACTCCTTGT
GAATGAATTACAATTTTATTGTAAGCGTTTCCTTGTAGTTTTCAACTTAAC
TCAAAAATTGTAATTTTTTATTAATAAATCAATGAACTTTTTATTAACATT
GAATTTGAATAGTGTGAAATTTGTCAAAAAAGGTTTATGTTTAAGTTAAAA
CCAACTTTAAAGAATTTAGAAAATTCAATTAAAGTGAATTTAACACATTTT
AAATAATGAAGGTAGATTAGAATCAAAATGTAAGAGGTATGTTAGTTGAAT
AAACCGTATTAACTTCTTTATATTATTAGTGAATTCTCAAAATACTTGTAA
AATTATATTGATTAATTACTTGTTCTTTATTATATTTAGAAACATCTCTTA
CAATAAACCACAACTTACATGAAATTTAAGCTAACAGTGGATTATTCTAAT
TAAATTGAAGGTCATTCAAAGTAAAAAGGGCATATTAGTATAATAGGAGGA
ACAGTAAGATGAAAGTGTATAAGTTGAATTATCAACACCATAAAGATATTG
TTGATGATAATGTATTAACGATGTTCGTAACAGCTGATAATCAAGATGAAG
TTGAAGCGTTTGCGAAAAAATTACATTACAAGATTGAACATTTATCTCCAT
TAACTAAAAAGGAATTTGAAGATGAGAAAGCGAAAGATTCACACTATAGAC
TTGAACACGTGGATCACTATTTAAATTAATTATTCAATCATATATCAGAAC
ATCCTGTGATCATTCGTTTGGTTACAGGTTTTTTACGCATGAGATTAAATG
TTAATGTCACTAATACAAGGAGTGTATAAATGGAAATAACAAATGTTAATC
ATATTTGTTTTTCAGTGAGTGATTTAAATACCTCTATACAATTTTATAAAG
ATATTTTACATGGTGACTTATTAGTATCAGATAGAACGACAGCATATTTAA
CTATTGGTCATACTTGGATTGCACTGAATCTAGAAAAAAAATATACCAAGGA
ATGAAATAAGTCATTCCTATACGCACGTTGCTTTCTCCATAGATGAAGAAG
ATTTTCAACAGTGGATTCAATGGCTTAAAGAGAATCAAGTAAATATTTTAA
AAGGGCGACCAAGAGACATTAAAGACAAAAAATCGATATATTTTACAGATC
TGGATGGGCATAAAATTGAATTACATACTGGAACATTAAAAGATAGAATGG
AATATTATAAATGTGAGAAGACGCATATGCAATTTTACGATGAGTTTTGAT
ATTTTAATTGATATGTAATTTTTTGAAAAAATTTATGATTCATTTTGTAAT
AAAATTACTATATACTAAAATTTGTAATTTTCAATTATTTTAATGAGGTGA
ATCAATGAATAAATTATCAAAGTACATTGCAATAGCTACATTAGCATCGAC
AGTTACAATTTCAGCACCAATTCATACATATGCATGTGAATCTCATACTAA
AGATAACCATAATCAAACGACACAACATCAAAATGACCCCAACCTTGGTGA
ACAAAATGTAATGGCTGTCTCATGGTATCAAAATTCTGCCGAAGCGAAGGC
ACTTTATCTACAGGGATACAATACTGCAAAATATCAGTTAGATGAACATAT
TAAAAAGGATAAAGGTAAGAAAAAAACTTGCTATAGCTTTAGACTTAGATGA
AACAGTTCTTGATAATTCACCATATCAAGGATATGCTTCTATGCACGATAC
GTCTTTTCCAGAAGGATGGCATGAATGGGTTGCTGCAGCGAAGGCAAAACC
TGTTTATGGCGCCAAATCATTCTTAAAATATGCTGACAAAAAAGGTATCGA
TATTTACTATATTTCGGACCGTGATAAGGAAAAAGATTTTAAAGCTACAAA
GGAAAATTTAAAAAATATTGGACTACCGCAAGCGAAAGATAATCATATTTT
ACTAAAAGGGAAAAACGATAAAAGTAAAGCATCGCGACGTCAACAAGTCGA
AAAAAATCATAAGTTAGTGATGTTATTTGGTGACAATTTGTTAGATTTTAC
AGATCCTAAAAAATCTACTGCTAAAGAACGTGAGAAACTCGTACAGAAGCA
TGCAAAAGATTTTGGCAAAAAGTATATTATTTTCCCAAATCCAATGTATGG
AAGTTGGGAATCGACACTTTATCATAATCAATACGAAATAAGTAAGAATGA
AAAGGATGAACTACGAAAGTCATCAATTAAACAATTTAATCCTAAAACGGG
TGAAGTAAAATAATTCCTGACATTAAATACTAAATTTCTATTAGTAAAAGA
AAGTTTTACACATAATTATTGTATTGTGTGCAGACACTTTAATAATTAATA
AAGTTATATTTACATGGAGTACTTTAATTCCGTGTTATACCATTGCGAGGA
```

TGAAACAATGTACGCGTTGTTCCATCCTTTTTTTTATAAATATTTTCAATC
GTATCTAGATTAATGTAAACTTATAATATCGAAAAGTTTACAATGTGCAAT
CGAATGATTTGAAAATGCTAATTTTTAAGAAAAAGGAGTTAAAGATTATGA
CATTAAACCTAGCTCAACGTGTGTTAAATCAAGAGTCATTAACAAAAGATG
AAGCAATATCTATTTTCGAAAATGCTGAAATTGATACATTTGATTTATTAA
ATGAAGCCTACACAGTGAGAAAACATTACTATGGTAAAAAAGTTAAGCTTA
ATATGATATTAAATGCTAAAAGTGGTATCTGTGCAGAAGATTGTGGGTACT
GTGGGCAATCTGTAAAAATGAAAGAAAAGCAACGTTATGCACTTGTTGAAC
AGGACCAAATTAAAGAAGGCGCTCAAGTGGCAACTGAAAATCAAATCGGTA
CATACTGTATTGTTATGAGTGGTAGAGGTCCTAGTAACAGAGAAGTCGATC
ATATTTGCGAAACAGTAGAAGATATTAAAAAGATACACCCACAACTAAAGA
TTTGTGCGTGCTTAGGATTAACGAAAGAAGAACAGGCTAAAAAATTAAAGG
CTGCTGGTGTCGATCGTTATAATCATAATTTAAATACGAGTGAGCGTTATC
ACGATGAAGTAGTAACTACACATACATATGAGGATAGAGTGAATACGGT

Sequence 3336
step.1000c01.cons.ok
TAACATCAGTGTTCAGTAAAAATTGGCTTCATTACTGGAGAAATGAATGTG
GTTATACTATAAAACCGGTTTCTGAGACAAATGACGTTCGGATATATGATC
AACAACAATTTATAATGTATGCGCATTTTGCTGATGAAGCTTATCAAAAAA
TTGATTACATTAATTATTTTGACACATCTAGAAGAAAAATAAAAAGAGAGT
TATACGATACACGCGGATTTTTAAGTTGTGTTAGAGTCTTATCAACGGACC
AAAAAATACAAGCGGAGTACTATCTTTCTCCTCAAGGAAACGTGAAAATTG
AAAAGTATTATGATATTAATTCAAATCATCCCTATGAGGCAAAAAAGATTA
TATTAAATCATTTAGGTAAAACGCACTTTTTAAACAATGAAACTGAGCTAG
GTGCTTTTTTTATTGAAACAATATATCAAAATAGTGACTTATTCTTTAGCG
ATCGTAACCTTGTTACATCACATATATTTAACATAGTTGCATCTTACATCC
CTGTTATTGCAGTGTTACATAGCACTCATGTTAAAGACGTTACTGATTTAG
CGCATTCACCAATCAAAAATGTTTATAAAGGGGTTTTTGAACATTTACAAA
GGTATAAAGCCATCATTGTTTCAACCCAACAACAAAAAGCTGACGTAATTG
AGAGAATCAGTGGAGTTATTCCAGTATATGCTATTCCAGTTGGCTATTCAT
CTATCGATATGAAAGATTACTCTAATGAAATAAATATGTTTCACCTAAGA
AAATTATTTCTGTCGCAAGATACTCACCTGAAAAGCAATTGATACAACAAA
TCGAACTTATTAATAAGCTTAAAGATTCATTTCCAAATATCGAACTACATA
TGTATGGTTTTGGTAAAGAAGAACAACATTTAAAAGAACGAATTCAAGAAT
TAGGATTAGAAAAGCACGTAATATTAAGAGGTTTTTTAAAAGATTTAACTG
ATGAATATCAAGATGCATATCTAAATTTAATAACAAGTAATATGGAAGGCT
TTTCCTTAGCATTATTAGAATGTGAGTCACATGGTGTACCTTCTATCAGTT
ATGATATTCAATATGGTCCAGGTGAATTGATACAGGATGGTAAGAATGGAT
ATTTAGTAGAAAAGAATAATCAACACATGCTATTTGAAAAAGTAAAATTAT
TGCTCAACAACCCTCAACTTCAACAGTCTTTTTCTCACCATTGTATTGAAA
CTGCACAAAAGTATTCTCGAACACAAATCATGTTACTGTGGAAAAATTTAC
TTCAACATTTTAACTAAAGCATAAAAATATAAGTCGTCAAATTTTGCTATG
AAGTCATTTATCATTAAATGGTTTCATAGTATTTTTTTGTATTAAATAAAA
GAATTTTTATTAATTCAATAAAAAGGAACCATGAATTCATTTAAAATTCAC
GGTTCCTCTATTAAATCATTGTCAGATTGTGAGTTGTATATGTTTCTCTAA
AAATCATACATTTAAGTCATTTTTAATAGAAATGACCTTGTTTGCATCATC
TTTTTCACAAGATATGATAAACGTGAAAGTTATATTATAGTTTTAATCAAC
ATCTTACATTTGTTGGACTAAATCTACTTTCTATTCAAATCTCATATCTGC
TTCTTTCTACAAATTTCATTTGAACATTTTAATTTTTATTTTTTCTATTTT
TGCGACGTTTCCCTAATAATAACGCTCCTAAACCTGCAAACAGAGTTCCAA
GTAACGTGCCTTTAGAGCCATAATCTTCATTAGCTCCTGTATCAGGTAATT
TATCTTTTGTACTCTTATCTGAGCTATTGCCTAAATCTGAGTCGTTGTCTG
AGTCTGAATCACTATCGGAATCCGAATCACTGCCTGAGTCTGAATCACTAT
CAGAATCTGAGTCACTATCCGAGTCTGAGTCACTATCCGAGTCTGAGTCGC
TATCGGAGTCTGAATCACTATCGGAATCTGAATCACTGTCTGAGTCTGAAT
CACTATCAGAATCTGAGTCACTGTCTGAGTCTGAATCACTATCGGAATCTG
AGTCGCTGTCTGAGTCTGAATCACTATCGGAATCTGAGTCGCTGTCTGAAT

```
CTGAATCACTGTCTGAATCTGAATCACTGTCTGAGTCTGAGTCGCTGTCTG
AATCCGAGTCGCTATCAGAATCCGAGTCGCTGTCTGAATCTGAGTCGCTGT
CTGAATCCGAGTCGCTATCGGAGTCTGAATCACTATCGGAATCTGAGTCGC
TGTCTGAGTCACTATCGGAATCCGATTCGTCATCATAGTATCCGTTATCTA
```
5
```
TACTAAAGTCATCATGATCAGTAATTGTTACATGAACTTCTTCCCCATCAG
CATCCTGTTCGTCATCATCACCAGAATCTGTTGTTGTTTGAGTCATACCTG
AAGGTTTATCAAAATGAACAATATAATTACCACTATTTAAATTATCAAATT
GATACTTTCCATTTTCATCGGTTGTAGTTGTACTAATGATATTTCCGTTTT
CATCTTTTAACGTCACTTTAACTCCAGAGATTCCTTTTTCATCATCACCTT
```
10
```
GAATACCATCTTTATTAGTGTCATACCATACATAGTTCCCTAAGCTGTATT
TAGGTGTTTGATGGGGAAAACAATTATGTTAAGTATTAAAAACTTAACCAA
GATTTATTCAGGGAATAAAAAAGCGGTAGATAATATTTCTTTGGATATTCA
ATCTGGTGAATTTATCGCTTTTATTGGGACGAGTGGTAGTGGTAAAACAAC
TGCACTACGCATGATTAATCGTATGATTGAGGCGACAGATGGACAGATTAT
```
15
```
GATGAATGGAAAAGATGTCCGTAATATGAATCCTGTTGAATTGCGGAGAAG
TATCGGTTATGTCATTCAACAAATTGGTTTGATGCCACATATGACTATTCG
AGAAAATATTGTTTTAGTACCTAAACTTTTAAAATGGTCTAAAGAGAAGAA
AGATGAAAAAGCTAAAGAACTTATTAAACTGGTAGATTTACCTGAAGAATA
TTTGGATCGTTATCCAGCTGAATTGTCAGGAGGGCAACAACAACGAATTGG
```
20
```
TGTTGTGCGCGCTTTAGCAGCTGAACAAGATATTATATTAATGGATGAACC
TTTCGGTGCATTAGATCCTATTACACGCGATACATTACAAGATTTAGTAAA
GGAATTACAACAAAAATTAGGAAAAACATTTATTTTTGTCACTCATGATAT
GGATGAGGCTATTAAATTAGCAGACAAAATATG
```

25

```
Sequence 3337
step.1000d09.cons.ok
AAACAGAAACGCGAGCTTCAATTAAAAACCTCATCACCTTCTTTAGGCTCAA
AGCCTAACTTAGAAGCATTTCCTTTAAACATCATTCCACTGATGACACTTT
```
30
```
CTTTGTCCTTGACATTAAAATATAAATGTCCGCTTGAATGCTTTTTAAAGT
TAGATAATTCTCCTTTAATAAGAACTGACTGTAGATGAGGATCTTGGTCAA
ATTTATATTTTATATATTTAGTTAGTGCAGAAACACTTAAATATTCAGTCA
TAGTTATCACTCATTTTAATCATTTATATTACTTAAAACACCATTAACAAA
TTTATAATGATCATCATCACTAAACTGTTTTGTGAGTTCTACAGCTTCATT
```
35
```
AACAACTACTTTTTTAGGTGTGTCGCTGTGCAATATTTCAAAAGTTGCCAT
TCTTAAAATAATACGATCTGATTTCAGTAAACGATCGATAGACCAGTCTTT
TAAATGGGGTTTAATTGTTTCGTCTAAAACGATTTGATGATCTTTGACTCC
AGTAACTAACCAGTATATAAAATCAAAGTCTAAATCAGAATGATCATCTTT
AATAAATTCAATTGCTTCTTGAATTGTTAAATCTGTCTCTTTTATTTCAAG
```
40
```
TTGAAATAAAGTTTGAAAAGCTTGTACTCTTGCATCTTTACGACTCATATT
TTACTCCTTCAAATGTATTTATAAAATTATAGGTACAACTTTCTAGTAAAG
TGCACCTATAATTTAAATTATTTATTTTTCTGCGACGATACTTCTGATGTG
AATATTAATTTGCTGTGGTTCAATAGCTGTCATTGTCGTAATTGAATTGAA
AATTGCTTCTTGAATTTGATTTGCAGTTTTAGAAATATTTACGCCATGTTT
```
45
```
TAAAGAACAAAATACATCAATGTATATTCCGTCTTCTTTAGCTTCGATTTT
TAAATCTCTGTTTAAATTTTTTCGACTAATCTTCTCTAGATTTGTTTTTTT
TAGTTCAGCAAAATGGCCTGTAATACCTTCTACTTCTGATGTCGCAATGGA
TGCGATAACAGATAATACTTCTGGTGCTATTTCAATTTTTCCTAAATTAGA
TTGAGAATAATCTGCTACATTGACCATTCTTTATCCTCCTATTGGTTATTA
```
50
```
TCGTCCATAATATTATACTTTTCTAGGAATTTTGTATTAAATTCCCCACTT
CTAAAAATATGATTATTTAGAAGTCTTAAGTGGAATGGAATCGTAGTGTCG
ATACCTAAAACAAGATATTCACTTAAAGCACGAATGCCTGTCATAATTGAT
TCTTCACGTGTAGGTTCGTGAACTATAAGTTTTGCCACCATGGAGTCATAG
TAAGGTGGTATCGTATAATTAGTATAACATGCTGATTCAATTCTCACTCCA
```
55
```
AAACCGCCTGGAGCAAGATATTGGGTAATCTTGCCTGGTGATGGCATAAAG
TTTTTGTAAGGATTTTCAGCATTGATTCGAAATTCAATAGCGTGACCGTTA
ATGGAAATATCTTCTTGTTTAAAAGGTAACGCCTCACCCATAGCAACTTTG
AGTTGTAATTTTACTAAATCTACTCCTGTTACCATTTCAGTTACTGGGTGT
TCAACTTGAATACGTGTATTCATTTCCATGAAATAAAATTGGTTATCATCT
```

```
AAATCATATATAAATTCAATTGTACCTGCGTTTTCATAATTTACAGCTTTT
GCGGCTCTAATTGCAGCATTACCCATTTCTTGGCGTTTATCTTCACTTAAA
ACTGGTGAGGGTGCTTCTTCAACGAGCTTTTGCATTCTTCTTTGAATTGTA
CAATCACGTTCACCTAAATGTATAACGTTTCCATAAGTATCGCCAATAATT
TGAATCTCTATATGTCTAAAGTTTTCTATAAATTTTTCTAAGTATAAACCA
CCATTTCCGAACGCGGTTTCAGCTTCTTGTTGTGTCATACGGTAACCAGTT
TCAAGTTCTTTCTCATCACGAGCAACCCGAATACCTTTTCCACCACCACCT
GCTGTGGCTTTGATGATAACTGGATAGCCGATTTTTTTAGCTATTTTTTTA
GCGTCATCTATACTTTGAATAAGTCCTTCACTTCCTGGTACTACAGGTACA
TTGGCTTTAATCATTTCAGCTTTAGCAATATCTTTTATTCCCATTTTCTGA
ATTGATTGATAGCTTGGCCCAATAAATTTTAATTGGCATGCTTCACACAAT
TCAGCAAAATCACCATTTTCAGCCAAAAAGCCGTATCCAGGATGAACACCA
TCACATCCTGTAGAAGTAGCGATAGATAAAATATTTGGTATATTTAAATAA
GAATCTTTTGATAAAGTAGGACCGACGCAGTATGCTTCATCAGCAATTTGA
GTATGTAATGCATCTTTGTCACCTTCTCCGTGTTACAAATTATGTTACAAG
TTTGCATGAAGGTTTATTTAACATCGCTGCAGCTGTAGGTGTTCATAGTCC
AACGGAGATTACTTCCGACCATATTATCTATAGACAATTAGATGGCACTAC
AACGTCCATTCAGGATTATAAACTTAAATTAATTTCTTAAATCACAACATT
GAAAAGGAGCGAGACAGAATGCTGTCTCACTCCTTTTCAATTTAATTATCT
TCATTCATTTTACTTTTATCTTTTAAACCATTACGTTTGAGTTGTTTAACA
ACCCTATCCATACTCACATGGTCACCCTTTTTTGCATCTGATGATATGAAA
CCTAAATTTTTCTTAAGCTGATCTATATCCGCATCTTTATAATCAATATCG
ATATGTTCGATGGCCTTTCGATCTTTAAAATCAACTTTATATGTGACCCCG
TCTATCCCTTTGAATGCTTGTTCATCATTTTTAAACATTTGTTTAGCTTCA
TCTTTATCGATGCCTAAATCATCATATTTAATTGTTGCTATAGTAGATTGT
TTGAGAACTTCATCATCTTTGTATGTCAAAGATGTAATCACTTGTTTACCG
TTGACTTCGCCTTCATAAGTTTTTGTGCGTTCTTTCCCACATGCCGAAAGA
GATACAGCACTTATAACTAATACTACTAAAAGCGATATGACTTTTTTCATA
GTAGTCGCCTTCTCTCATCTTTAATCATGTTAATTATATCATCTTTAATTT
GAATATGTTACTTATATGAAGATAAATATCTTTTTATTTCGAAAATTTAAA
CATCTATAAACATGGAATTCAGAATTTCACTACTTAGAATGTTATACTACT
TTTAAAGACTTATTAATGGAGGAATGACTATGCCAGAATCTAAAGCGTTAC
TTAAATCTTTAACAGATGTTAATGGTATTTCTGGACATGAAATGCAAGTTA
AATCT
```

```
Sequence 3338
step.1000f04.cons.ok
TAATATCTTTTTTATTTGGTAACCAAACGTTTTCAGCTTGAGTAAATGGAT
AAATTGTATCTGATGCGGCTACTCGAGCTATTGGAGCTTCTAATGAAAGAA
TTGCTCGCTCTGCTAATTCTGCTGCCACTTGTGCACCCACACCAGCTTGAC
GTTGTGCTTCTTGTACAACTACAGCACGTCCAGTTTTCTCAACTGATGCTA
CTAAAGTATCTATATCAATTGGTTGTACAGTACGTAAGTCAATAACTTCAA
CTGAATAACCATCTTTTTCTAACTCTTCAGCAGCTTTTAGTGATTCTTGTA
CCATTGCCCCGTAAGATATTAGAGTAATATCATTACCTTCTTTTTTAACAT
TGGCTTTTCCAATGTCAATTTTGTATTCTTCTTCAGGAACCTCTTCACGGA
AAGAACGATATAATTTCATATGTTCTAGATATACAACTGGATCATTACTTT
GAATAGAAGAAATTAATAATCCTTTAGCATCATAAGGACCTGATGGAATAA
CTACTTTCAAACCAGGTGATTGAGCTAAGATACCTTCTAAATTATCAGCAT
GCAACTCTGGAGTGTGGACGCCACCACCAAAAGGTGTACGAATTGTAACAG
GCGCTGGTTTAGTTCCACCTGAACGGAAACGAGTACGAGCAATTTGACCAG
CTACTTCGTCAAATACTTCATAAACGAATCCTAAGAATTGAATTTCCATAA
CAGGACGGAAGCCAGTCACTGCTAAGCCTAGTGCAAGCCCACCAATTCCAG
ACTCTGCTAATGGTGTATCAAATACTCGATCTTCGCCAAATTCTTTTTTGTA
AACCTTCAGTAACACGGAATACACCACCGTTAACACCAACGTCTTCACCGA
AAACTAAAACGTCTTCGTCTCTTTTGAGTTCACTTTTAAGCGCATCGTTAA
TCGCTTGAACCATTGTCATTTGTGCCATGGCTTACTTCGACTCCTTCTCTT
TGTAAATTTCATATTGTTCTGCTAAATTTTGAGGCATTTCTTCATACATGA
TATCCATTAGAGAAGTAACAGTTTGTTTTTCTGTATTGTCAGCCTCTTTAA
```

```
TAGCTGCTTTTATTTCAGATTTTGCACGTTCAACCACTTCATTTTCTTTGT
CTTCATTCCAAAGACCTTTAGCTTCTAAATATTTTCTGAAACGTACTAATG
GGTCTTTTTTCTCCCATTCAGCATCTTCATCTGAAGTTCTATAACGAGTAG
GATCATCACCAGCCATAGTATGTGGTCCATAACGATAAGTTAAAGTTTCGA
TAACAGTAGGACCTTCTCCTGCTACTGCACGTTCACGTGCTTCTAATGTTG
CTTGATAAACAGCTAAAGCATCCATACCATCAACTTGAATTCCAGGGATAC
CAACTGAAATAGCCTTTTGTGCTAATGTTTCTGCAGCTGTTTGTTTACTAC
GTGGTGTAGAGATGGCATAGTTATTGTTTTGAATTACAAAAATTGCAGGTG
CTTTGTATGCAGATGCAAAGTTAATTCCTTCATAGAAGTCACCTTGTGATG
AACCACCATCACCTGTATAAGTAATTGCGACTGCATTTTTGCCACGTTTTT
TAAGTCCAAACGCTACACCGGCAGTTTGAATATATTGTGCACCGATAATAA
TTTGAGGGCTAAGTGCATTAACTCCCTCAGGGAATTGGTTACCTTTGAAGT
GTCCTCTTGAGAATAAGAATGCGTCTGTAAGAGGTAAGCCATGCCAAATAA
TCTGAGGCACATCACGATAACCAGGTAAAATGAAGTCTTCACTTTCTAAAG
CATACTGAGATGCTAATTGTGAAGCTTCTTGTCCTGCTGTTGGTGCATAGA
AACCTAAACGTCCTTGTCTATTTAACGAAATAGAACGTTGATCAAGAATTC
TAGTCCATACCATTCTTTCCATTAATTCCACTAATTGTTCGTCTGTTAAAT
CAGGTACTAAGTCTTCATTAACGACATTTCCGTCTTCATCCAAAATTTGAA
CCATTTCAAATTTCGATTGAGTCTCATTTAAAACTTTAACTGCATCGAATT
GGGCTTGTAACTTAGGAGCCATTCAATTCACCATACCTTTCCCATGTAAAT
AGAAATTCATTTTATCCTCCTTAATTGTATCACAAATAAAAATCTCTGTTA
AACGATTTCATAAAGTTCTGTTCCACATAAAAATAATACAGTTACATGACC
TGTTGTATTAAAAATCACGTAACTGTATTAACCCATTATTTATTTTAATTG
GTCTACATCTTGTTTTTCTCTTTTTACTTTCTCAATGGCTTTTGAGTAAGC
ATTAAATTTATTATTCATTTCTTTATATGCTTTAGAAAGATCTTTCGATTT
TCCGTCTACTTCCGATTGTGTTGCATTATCTTCATTCAAATAAGAAAACAG
TTCTTTTTCCTTATTAAGTGCTTTTTTTGTAAGCCTTTGCATAAGCATCATG
TGATTTATATTTTTCTTTAATAGCACTATCAAGTTGTTCAACTTCTTTTTT
CTTTGCTTTGTTTTCTACATGTTCAAGATATTGCTTGGCTTGTTTAAATGC
TTTTTCAGAATTATCTAGAGCCTTCTCTTCTTTTTCAAATTCTTTTTTGTCT
TTGTTTTACATTTTCAACTATATCTTCAGCTGCTTTTTTACGTGTGCTTTG
ATCTTTATCATTTACCTTTTTAAATAATTTTTGCTTTTTCTCTTCTAATTC
GTTTATTTTTTTTACTTACTTGATTAACGGTTTTCTCTTGGTCCATAGCTTT
TTGTACTTGATCATCATATTTTTTAATTTCTTTTTTATCTATAGTACATCC
AGCTAATAACACTGCTGACGTTAGGACAATTGCTACAGTCTTTTTAAAATT
CATGAACTGTTACCTCCTCAAAGTTGTTCGCTTGTAATATTAAATAATAAA
TTATAATAAATCAATATAAAGCTATCATTCTATATTCTAGTGCAAACATTT
AAAGCAATAAAATCTTTGAGTATCTATTAATGTACTTGTGTCTATCTTATC
TGATATATTATTTAATAAGGAAGGTGTATGACATGATAACAATGAAAGATA
TTATAAGAGATGGTCATCCAACACTTCGTGAAAAAGCGAAAGAATTAAGCT
TCCCACTTTCTAACAATGATAAAGAAACATTGCGCGCAATGCGTGAATTTC
TAATCAATAGTCAGGATGAAGAAACCGCAAAACGTTATGGTTTACGTTCTG
GCGTAGGTTTAGCTGCTCCACAAATTAATGAACCAAAACGTATGATTGCTG
TCTACTTACCTGATGATGGAAACGGTAAATCGTATGATTATATGCTCGTAA
ATCCTAAAATAATGAGTTACAGTGTACAAGAAGCTTATTTACCAACTGGCG
AAGGTTGTCTAAGTGTTGATGAAAACATCCCAGGTTTAGTGCATCGTCATC
ATAGAGTCACTATTAAAGCTCAAGATATTGATGGAAATGATGTTAAATTAC
GTCTCAAAGGCTATCCTGCAATTGTATTTCAACACGAAATTGATCATCTAA
ATGGCATTATGTTTTATGATTATATTGATGCCAATGAACCTCTAAAACCAC
ATGAAGAGGCCGTAGAAGTCTAAATCATGAATTATATTAAAACACTTCATT
AAGTTGTATTTTATACTTACTTAATGAAGTGTTTTTTTAAAAGTAATTCAA
ATGTCATTAGGATTAATTT
```

Sequence 3339
step.1000f07.cons.ok
```
TCAATATTATTTTTCCTTCAATATCATTTTATTAATATAATTAGGTAGAAG
TCTACGCTGTCTCTAAGTGGTAGATGAATGATCTCTATTTTTTATTATGTT
ATACATTTATATTATTTTCTGTAAGGTTCTGATTCTAAAAGATATTTATAA
```

```
        TAGGAGTCTTTCTTTTGTAAATCTGTTTCCCCCCACACCTTATATTCTTTA
        TCATCTGTCCATGGTGCTGGTGTCACGTTATACTCATACTCACTATATGGT
        TGATCTTTAAAAACAATATCTTTTTCAGCATATCCCATTATAATATTAAAA
        CTCATTTTTTCATTTTTTATTTTTTTATCCCAACCCTTACGTTCTATAACG
  5     TGATTGATGATTTTCTCACTCTTGTGTTTCTGAATATATTGTAGTATAAAG
        AAAATAAAGTTTATAAAATAGAGCTTGATGATATGAATGACAATAAATAGA
        CAATAAATAACCGCATCATAATCTAAAATGATTATAATACGGTAAATTTTG
        ATATTATTTTATAACTCAATTGTAGCTGGCGTATCTAGTTCATCATCATTC
        ATTAATTTAATTGAACGAGGTATAACTTTGAGTACACATAATTCAGGATCT
  10    TCTTTTGAATTGAAAAATGTTTTGTCTTGAGTTTCCCATAACCAATCAATA
        ACTTTTTGATTTTTGACAACTTCTATAGTGGCATCTATTTCAACAAAGCTA
        TTGTTTGTTGTTTCGTTATAGCCTAATAAAATATGAGCATCAGGATTATTT
        TCTATTTCTTCGACTTTTAGTGAATTGATATTCGTTTTTGTATATAAGTTT
        AAGTCATCGTTGTAAAAGACCATATATCTGCTATTAGGTTTATTATGATGT
  15    GCTGTTGATAGGACACCAATTTTTGAAGAATTTAATACTTTTTCTATTGCT
        TTTGTCACTTGTTGTTTATTCAAAATTATCATCTCCTTATTATTGTATATA
        TCCTTATCACACATAATTTAAACGAAACAACCTTATGATGATTTTTAAGAA
        GGTTTAAAACAAATATATTTGACTTCATCGTGTTTTAAGGTGACTTGGAAA
        GGAAGATGATGGTCATGTTGTATAGCTAATTCTGTTTTTGGTTGATTTGAT
  20    TTATCAAGTTCATCAAGAATACTACGTTCAATATAAAATTGTTGTTTAGTT
        TGTGGTAGAAGGTTTTGAATTGCGCCATGCTCATGATTTAACGTTTTAATA
        ATAATTAATGAGTTGGTTGATAATGTATTTTTAAAAACGTAGCGCTGTTGA
        CTATCAGATAGATAACGATCATTTATTTTATTAAATAATAAAAAGTGATAG
        TTACCATCTTTGCGACTCATTAAAAAATTGTTAGCTATCTCTAAAGGTTCA
  25    TTTAAAAATGGTTCGATTAATTTATAGAGGTGCATGAGAGGTAGGGGACTA
        CCATAACGATTAAATAAGGCAATAAAAGGACTCTCATCTTCTATCAGTTGA
        TAACTTACACCACCCCCGTAACGCATTAATGCGACTAAATGACAAACAAGG
        TCAGATAATTGAAGTGGATAAGTACCATTGTTTGTATATTTAAAACACTTT
        AAAGATAAATTATTGTAAACAAATTTAGTTGAAGGTATACCGGAATCTAAT
  30    ATAAATTGTTTATAATTTTCAAAGTGTGTGGACGAATGTAAGTACGATTTA
        CGTTTGGTAATTAATGAGTTTTTGGTTTCTAGATTTTCAATGTCTATAAAA
        TAGTAATCGAAATTTAATCGCTTCATCATATCATAAGTTCGATCAATAGAA
        GAATTTTGGTGTAGTAATCCTTCAACAGTTATACTATAACGGATAGCTTTA
        TTTTTATTTTTAATTTTTAAATGGCATAAATGTATATCGTTCACTGACATA
  35    TTCTTAGTTTCAAAGACTAACATAAATTTAACCATTTTTTTGTTCATTTTA
        TAATCTTGGTGGCTATTAAGAAATAATAAAATTATTTCTTTGATAGATTCA
        AATTCATTTGTAGATGTTAATCTCATGGCTAAACCTATATTTTTTTCGAAT
        AATTTTTCAAAGCATCGATTTAATAAATTAAAGTTCACCTCTCGCGTCGAA
        ATATCAGTGATATCATTGATAAAAATTACAGCTTGTGGCAGGGAGGTAAAA
  40    TCAATGTTGTAATATTCATTAAAAACAAATTGAAATAATTCGTTGAAATTT
        TGAAAACGTATAAACGCTTTAGAGTTTTTAGTCTGATCTTGAGGATAAAAT
        TCATTGATATCTCTTTCAGTCGTTTCAGTAGCCAAATGATCACTAAACTCA
        AATTGATTAATTAAATCAATATATTGTGAAAAGTCAGTATCACTAAAACTT
        ATCGAAGGTAAGGAGTCTAACTTAGAGATAATCGTTCGGTATTGTTTTGGG
  45    CTACAACCTAAATAATTTTTAAATTGATTTGTAAAATTGGTATGACTACTA
        AATCCAGATTGTTCTGAAATAGCGTTAATTGAGTCCTCAGATGAAAGTAGC
        CTTTTTATAGAGTCAATCACTTTTAAGCTTGTAAAATAATCTTTAAAATTC
        ATATTAAGATATCTTGCAAATAGGTTAGAACAATAAGATTCAGATATATTA
        CAATGTTGAGCTACTTCTCGTAGTGAAAGTGACTGTGATACTTTTGAATGA
  50    ATAAATGTTAATCCTTCAGTAAACACTGAATGATTAACTGCAATTTGAGGA
        ATATATTTTTTCTTATATCGAATGACTGCTTCTTTTAGTAGCGTTTGTATT
        ATTTTAGAGATGGATTGCTCATCTTGATTTTCTCCTTTTATTAAATGTTGA
        ATAGCTTGTAAAATAATTGTTTTAACAAAACTGCTTGATTGTAATAAATGT
        CTGTCAAAATAACATTTAAAAAAATTTATTATCTTTACTATAAAAATAAATA
  55    ACTGGAATAGAGAGTAATACGAGATTTGAAGCACTATTAATAAAATAGATG
        TCAGATTGATTAATGATAGCTATATGATTACCTAAATCTTTTGTTTCACCA
        TTAATAGTGATTTTTAAGTCATTTGTAAGTGAAAACATCAGAATAATTCTA
        TTAATATTGCGTTCAGGGGGAGAATGATATTCTTTATATATATTTAAAAAAA
        CATGTCATTGTATATTACATTAGCTCCTTGAATGAATTGAATAACTAAATT
```

TAAATGTGACCGTGT

Sequence 3340
step.1000f11.cons.ok
CCATAACCATGATCAATAACTAGTGGATAATATTTAATCCGTCCAGAATGT
GCAAAATAACGTTCATCTTCAACTATCAAATGATTAACCTTACTCATTGAG
TTCCCTCCCAATAGTGACTCTTTTTAATATAGCTATCTACATATGAAGCTA
TAAAAACTATCAACTGCTATCGTTTTAAAATAAATCAATATTAATAGGATA
GTAATCAATATACTAATAATTTTCAGATTTGTAAATATTTTCTGATAATTT
AGAATACTTACGTTCCGAATCTTTTGTGATAAAGAAATATATGTCATTGAG
GCTTAGAATATATGTTTAAAGGGAATTTAGAATAAAATATTTTACGGAGGA
GATTAATATGCATTGGTTAAAAATATTTTATCATTTATTATGCGCAACCAC
GATTAGCGTGATATTACTTATTATAACTATATTAATGGATGCGTTACTACA
AAACACACACTTAACTCAGTTATTACTCAATATTGATTTTTTAATCAACCC
TGATGAAGTGCCAACAATTATTGAAGTACTGATTCATTTAAGTATTGGAAT
ATTGATTTATCTCGCCTTTTTAATTATCTATCATTATTCAAAATCCTTGTA
TCATCTAGCATACTTACCTTTAGTATTGATATTTACTTTGATGTATCCACT
TCTCGTTTTTCTTGCGCAACGTCCATTTTTTTCCTTTAGTTGGAACGAATT
TGCATGGTGGTTAGTTGCACATCTTTTTTTTCATCATTTTAATGGCGACTTG
TCTACCTATCATTTCGAAAAAAAATTTTATGATTTTAAAACTCATATCTCAA
TGTCTTTTACACATTATAAAATAAATATTGGCAGTAGAAACATCATAGATA
ATTAGACTATGATATCTCTACTGCCTTATTAGTTATTCTGGCTCTCCATCT
TCATCATACTTCGTTACTTCTCCATCTTTATCGATAATATAAGATCCTTCT
AAATGGCCTTCTTTATCTGTAAATGAGAACCCCCAACTACCATCGCTATCT
TTTTCTGGTTCTTTGTAAATATATGTGTCTGTATCTAATTGATGCCCTTCG
TAGTCTTCAACTATATCGATAACATTACTTCGTGTCACCGTTGCTTCATCA
GAACTACTTGATTGATCTTCAGATTCAGAAGTATCATCTTCATCTGTTAAA
TTATCACCAACTGAAAATTCAGTCCTCATCATCTGACTTATCTTATTGATT
TCACTTTCACTAGCGTTGTATAACTTAACTGTACGAGCATCATCTATAATG
CGTTGAGATTCTCTTTTAGAGTAATCAGCGTCACGCCAACTCCCTTGGAAA
TGTGATGGTACACTATATATTGTAATTGTGCCĀTCTCCATTACTTTTATAA
TAAACTGAACCTGCTGCAGTAACACTTGCTGTTAACAATTCTGTTCCTTCA
GGTAGTTTGGCTGATTCGTCTGGGTGATAGGGATCTAAAACTTTATTAGAA
ATATTTTGATGTACAATTTCTAAATCGTCAAAAGGAAGGTCACTTTCACCT
CTATAACTATCTAAAGCAGTTAACCAAACTTGAGCATAGTATCTATTCGAA
TGGGTATCTTGTTGTGTGTCCTGCTGATTATCTTCTGTTTGATTTCTACTT
ATACTTGATTTTGAATCTCTGTTACTTTTATATCCTGATTGTTGATCATCA
TTTATCGTTGAGCTTCCATTATTATCACTATCACCAAGACCACAACCAGCT
AATATTAATATTGTAAACGCACATATCACTGCTAATTTCTTCAATTCAAAT
ACTCCCTCACATTCTAATATATTTAACTCAACAAAATAAACCACTTTATAA
AATAATTATAAACTAACTATTTTTAAAAATGTACAACAATTTAACAAATTC
TTAAAGATTCGCTATTTACGAAGCAATACTTGCCGTCTGTTTTACTGCTAT
ATCATTAGCAATGGTTTATGATTATTAAGATTTTTAACGTATCTTAAAATT
TGTATTTGAATTCTTTTAAACTACTATCATCTCGCAATCAAGGTGCTAAAC
TATAATCGTAAAGTAATTTAAATACATCATATATCATCAATTCGAAATGGG
GTTATCAATGTGAAAAAGATTTCTCTAATTGCAACGACTGTATTAACAGGA
TTATTATTATTTCCAAGTGTCAATGACACAACGACACATGCAGCCGAAGTA
ACATCTCATGATGCAAAAGTGCTTTAAATGAAGCTCAGCAGTCATTGTCAC
AATATGACAGACAACTCAATGACAGCATGAATGCATCATTTGATGGTAAAA
TTAACATTAAAAATGATTCAGATGTAGGCGAAGGGCAACCTATTTTGCAAT
TAATTTCTTCAAATCCTCAAATTAACGCAACTATCACAGAGTTTGATATTA
ATAAAATTAAAGAAGGCGATGAAGTAAATGTCACTGTAAATAGCACAGGTA
AAAAAGGAAAAGGAAAAATTCTTAAAATAGATGAACTTCCTACAAGCTATG
ATACAAGTGACGATAGTACAGCATCATCGGCACAAGCAGGGGCACAAGGTG
ATAGTGAAGAAGGAACTGAAATGACGACATCTAATCCTACAATTAATCAGC
CAACAGGTGGTAAAAGTGGCGAAACATCAAAATATAAAGTTATCATTGGTG
ATTTAGATATACCCGTGAGATCAGGCTTCTCTATGGATGCTAAAATCCCTC
TTAAAACTAAAAAGCTACCAAATAACGTGTTAACAAAAGATAATAACGTAT

TTGTCGTTGATAAAAATAATAAAGTTCACAAACGTGAAATTAAAATTGAAC
GTAATAATGGTGAAATCATTGTGAAAAAAGGATTGAAATCTGGCGATAAAG
TCCTTAAAAGTCCAAAAGGTAATTTAAATGATGGAGAAAAAGTAGAGGTGT
CATCATGATTGAATTGAGAAAAATCAATCGACACTTTAAAAATGGAAATGA
ATCAAATCATATTTTGAAAGACATAGATATTCATATTGATGAAGGTGAATT
TATTGCTATCATGGGTCCGTCTGGTTCAGGTAAAAGTACGTTAATTAATAT
CTTGGGATTTATTGATCGTGGATATGAAGGAGAATACTTTTTTAACAATGA
GAATTATCAAAAAAGCTCAGATAATAAGCTCGCAGAAATTCGAAATCATAC
TGTAGGTTTCGTCTTTCAGAATTTTAAATTGATTCAAAATAACACTATTTT
AGAAAATGTAAGTATTCCGCTAATTTATAATGGTTTGAGTAATAAAGCGAG
AAAAAGCAAAGTTTTAGACGGACTTCATGACGTCGGTCTCAAAGGTAAAGA
GAATCTCTTACCAAATAAATTGTCAGGAGGGCAGCAACAACGTGTAGCTAT
TGCCCGGGCAATTATTAATGACCCTAAGTTTATTATTGCCGATGAACCTAC
TGGCGCTCTTGATTCGAAGACATCTCAAGATATTATGGAGCTTTTCGTAAA
ACTGAACAAAGAACAGAATACAACTATTATTATGGTAACGCATGATCGTGA
AGTGGCGGAAAAAGCTGATCGTATTATCCATATTTTGGATGGGCGTGTGCA
AGAGGAAGAGGTGTTGACACGTGAATAACTTTTCAAATGTCATATCAGTTG
CTATTCGTTCCATTTTAAAAAACAAACG

Sequence 3341
step.1000f12.cons.ok
AAAATTAGGTCCAGAAAGTTTCCCATCTTTCGAAATATCGGTATAAATCAC
ACCACCCAAAGGTAAATGCTCAATTTTGGCAACATAATCAAATAAATTGAG
TTTAGCATCCTCTTTCCATCCATTAATCTTTATTTTCTCTCCAAAAGCATC
TACGGATAAGTAGAGTTTATTTGGAAATTGATGTGTCATATGTGTTAACCA
CTCTATATCTTGGATACCTTTTGTACCTACAATACAATAGTCTATTCCTGA
ATGAATATAATTTTCAATTGTTTGTTTTGAACGAATGCCGCCACCCACTTC
TATAGGCTTAGTGGTCACCTTTCTTAAGGAACGGATGTAGTCGAATTCTTT
TACCTCTTTTGCTTTAGCCCCAATTAAATCTACAATATGTATACGTTTCAC
ACATTTAAATTGACTATAAAATCGGATACTGTCTTCTACAGATTTTTCCAT
TTTTTCTTTTGAATCATATTTGCCTTCTGTTAATCGAACACTTGTTGAATT
AATTAAATCAATAGCTGGCCATAAATCAATCATTAATAAACCCTCCTTTAA
GCGCTTGATTTAGAATCTCTAATCCATACGTTCCACTTTTTTCAGGATGAA
ACTGGATACCTATATAATTTCGGTATTGAATGACTCCCGGAATCTTTGTAC
CATAGTCAGCATAAGCTACTACATATTCTGACATTTCTGCTTGATATGAAT
GAACAAAATACACATCACTTTGCAGTAAGGGATGTGTACTCTTTAATTCAT
TCCAACCCAAATGAGGAATAGGATGAGATGATTGGATTGGCACTATATTTC
CCGGGACAAGTTCCAATCCACTAACGTCACCTTCTGCGCTATGTTGAAAAA
GTAATTGCATACCTAAACATATTCCAATTATCGGTTTATCATGTATATTTT
TAAGCATATCTTTAATGCTTTTTTTCTTCTATAGAATGCATCGCATCCTGAA
AATGTCCAACACCTGGAAGTACGATAGCTTCAGCTTTTTGCACGTCTTTAT
CATCACATGTCAAAATCACATCATATCCTAAATGTTGAATTGCTCGAGTGA
CGTTACTTATGTTTCCCAATCCATAATCAATAATCGCAATCATTCAATTAC
TCCTTTAGACGATGGAATACGTCCATCTTCATTTTGTGCAAGAGAAATCTT
TAATGCTCTTGCAAAAGATTTAAATATTGCCTCAATCTCATGATGTGTATT
TCCACCTCTTAATAAGTCAATGTGAACGGTTAATCGCGCATTAATTATCAA
TGCTCTAAAAAATTCTTCAACTAGTTCAGTGTCAAAAGTTCCTACCTTTTG
AGCGCTCAACTTGCTATTAAATGAGAAATATGGACGACCACTAATGTCCAC
TACTGTTCGAGCAAGCGCCTCATCCATGGGTACATATGAGCAACCATATCT
TGTAAAACTTTGTTGAGTCTTTATTAATTCAAGAAGTAATTGTCCAATAAC
TATACCTATATCTTCAGTTATATGATGATCATCAACATACGTATCTCCAGT
GGCCTCAATAGATAAAGTTAATCCACTATGAAAAGTAAATAGCGTTAACAT
ATGATCTAAAAATCCTACACCAGTATTAATATGAGATTGTGTTCCGTTATT
AGCCAGTGAAATATTAAGCTGAGTTTCTTCAGTGTTTCTTTTGATTTGATA
GTTCATAACATACGCTCCATTTCTTCACTATTTTTCCAGTAGAAGAAGTTG
TTCTTCTGTAGCAATAGAGTAACGAACATACGCTTTCATTTTTATTAACAG
ATATCGTAATAACCTATCTTTATCGGCGTACGTGAATCGATTTTTTCATGAT
TAAATAAAGATTCTATATGTGCAATATGCTCAGCTGATTCTGCCACTTCGT

```
     TAAATGTTTTTTGTGATAAATGAATGACTGAATGTCGTGTCATAAAGTCAT
     TCACAGAAAGCCCATTTGTAAATCTAGCTGTTTGATTAGTAGGAAGTACAT
     GACTTGGACCTGCCACATAGTCCCCAATAACTTCAGGAGAAAAATGACCTA
     AAAAAAGTGCGCCTACATATTTAACTTTATCTATATACATTTCAGGTGCTC
5    GAGTTTGAATAGATGCATGTTCAGGTGCAATCGTGTTCATAATTAAACATG
     CTTCTTCAGTATCTTGAGCTAATATTAGATAATGATGGTTAGCAATACTTT
     GTGATATGATATCTTGCCGTTCAACATACTGAAGTTTCTCTTGTATTATAG
     TGTTCAATTGATTAAGTACCTTTTCATTTTCACTAATCACATAAGTACAAG
     CCATTTCATCATGTTCTGCTTGTGCAAATACGTCATAAGCGATTGCGTCTA
10   AGTCAGCACTTTCGTCTATAATCAAGGCTATTTCTGTCGGTTCTGCGATTT
     GGTCTATGCCTACTTGACCGAATACAAACTTTTTGGCATAAGCAACATATT
     GATTCCCTGGACCTACGATTTGTCGACTTTTTTTATAGTTTCCGTGCCAT
     AAGTTAGCGCCGCAATACTTTGTGCTCCACCGACTTGATAAACATGATGAA
     CGCCTGTAATGTAACAAGCGGCTAACACCTCTTGACATATACCGCTATTTT
15   GAGGTGGGGTAACAACAGTAATCTCATTAACACCTGCTACTTGAGCAAGTG
     TTGCAGTCATTAATACTGTAGACGGATAGCTAGCCTTACCTCCCGGCACAT
     AAATACCTACACGTTCGATAGGATGGTATCGTTCATAACATTCAGTTTGTT
     GAGATGATTCCTGTTTTACCTTAATATTTTCTTGGTACACTTTAATTCTCT
     GATAGCTTTGCTCTAATGCATCTCGTGTTTCATTATCTAGCATGTCGTATG
20   CATTTTTTAGTTGGCTTTGCTCTAATTCAAGCTTCTCAGTTTCCACTTGAT
     CAAATTGTAAATTATAATTTTTTAAAGCTTTATCTCCCTGTAATTTAACTT
     CTTCACAAATCTGACTCACTATCTCGTACAAAGATTCATCTAGAGATTCAA
     CATTATTAAATTCTTTTAAAAATTGTTGTGCGCTAAGCATAGTTAATAGAC
     ACTCCTAACTGCTTGATTAGTCTCTCTATTTCAGATGATTGCTTAAAATAT
25   GATTCTTTATTTGTAATTAACTTAGCGTTAATTTCACTGATATGC


     Sequence 3342
     step.1000g04.cons.ok
30   ACACCATCCCTATCATATATACGGCAGTAAAGATGAAGACCCAAACAACAT
     CTAAGAAGTGCCAGTATAAACTTACTATAAATAATTTAGGAGCATTGTATG
     AATCCAAACCACGAGTGCCGATTTGAATTAACAAACAAATAACCCAAACAA
     TACCTAATGATACGTGTGCACCGTGCGTACCTAGTAGTATAAAGAAACTAG
     ACCAGAAGGAGCCAATAGTTGGGTTAACACCTTCAGAAGCATAGTGTGCGA
35   ATTCGTAAATTTCGAAACCTACGAATACAAGACCTAGGATAACTGTGATAA
     TCATCCAAAACATCATTAAGTTTTGTTTTTCTTGTCGCATGTAATAAATTG
     CAATACCACAAGTATAAGAACTAATTAATAATGCAAAAGTCATTATTAAAA
     TCAAATGTAATTCGAATAAGTCGGTAGTTAATTTACCGCCATATCCGCCAC
     CATGTTGTAACGTTAATAACGTTGCAAATAGGGTACCGAATAACGCAAATT
40   CAGCTGTAAGGAAAATCCAAAAGCCAAGTTTATTTAATTCGCCTTCGTGTG
     TACGAGAATCAATAGTATTTGCATCATGACTCATGACTTACAGCCTCCCTT
     TCTTTAATTCGAGCTTCTCTTAAACGAGCTTCAGTTTCTGCAACTTCTGAA
     GCAGGGATGTGGTAACCATGATCAATTTGGAAACTTCTCCAAATCATAGTA
     ATGAAGATACCTGCTAAACAGATAAGTGCTGGAACAATAGATTCGAAGATT
45   AAGAAGAAACCACCAATAGTCATAAATATACCCATCAGAATCCTACTGGA
     GTATTGTTTGGCATATGAATATCTTTGTAGTTATGGTTGTCTAAATAATGA
     CGACCATGTTCTTTCATATCAACGAATGTATCGTAGTCATTCCAATCAGGA
     GTGATAGCAAAGTTGTATTTAGGTGGAATAGCTGATGCTGTAGACCATTCT
     AAAGTACGACCAAGTCCATCCCAGTTATCTCCAGTAGCTTCACGTGGAGCT
50   TTGATATGACTATAAACGATACTTGCAACTAGGAATAAGAATCCAATTGCC
     ATCAATACTGCACCGATAGTTGAGATGAAGTTTAGTAACCACCAACCATCA
     GAAGGCATGTAAGTGTATAGACGACGTGGCATACCATCTAAACCTAGAATG
     AATTGTGGTAAGAAACAAACGTTAAATCCGATCATGAAGAACCAGAAGCAC
     CATTTGTTTAATGTTTCATTTAACTTGTAGCCCATCATTTTTGGATACCAG
55   AAGATTAAACCAGCTAAGCAGGCAAATACTACACCAGTAACCAATGTATAG
     TGGAAGTGAGCTACTAAGAAATAAGTGTTGTGATATTGATAGTCAGCTGAT
     GCCATTGCAAGCATTACACCAGTAACCCCTCCTAATAAGAAGTTAGGGATG
     AATGCTAATGAGAATAGCATAGGTGACTCAAATGTAATTCTACCTTTGTAT
     AATGTGAGCAACCAGTTAAATAGTTTAACTCCGGTTGGAACACCGATTAAC
```

ATTGTTGAGATAGAGAAGAATGAGTTAATTAACGCACCATTACCCATAGTG
AAGAAATGGTGAACCCAAACTAAGAAACTTAAGAATGCGATACCTGCAGTT
GCCCAAATCATACTTTGATGACCGAATAAACGTTTACGGGCAAAAGTAGGG
ATGATTTCTGAGTACATACCGAATGCTGGCAAAATAACGATATAAACTTCA
5   GGGTGCCCCCATACCCAGAAGAAGTTTGCCCAAAGCATTGGCATACCGCCA
TTTGCTACAGTGAAGAACTGAGTTCCAAAAATTCTATCAGCAGTCATTAAA
GCAAGTGCTACAGTGAACACTGGGAATGCTAATATAACGATTAATGTTGTA
ATGAATGTTGTTACACTGAACATTGGCATTTGCATAAACTTCATTGTTGGA
GTTTTACATCTTAGAATCGTAACAAAGAAGTTGATACCAGTCATTAACGAT
10 CCGATACCAGATATTTGAATTGCAATTAAATAATAGTTAACACCTGGACCT
GGACTGAATTCACCAGCAAGTGGTGCGTAGTTAGTCCAACCAGCAGCTGGT
GATCCACCTACAATAAATGATAGGTTGAATAAAATCATACCAGCGAAGAAT
AGCCAGAAACTAACGTTATTCATTACAGGGAAGGCAACATCGCGTGCACCA
AGTTGTAATGGAATAACAACATTCCATAAACCAAAGATAAATGGCATAGCC
15 ATAAATATAATCATAATTACGCCGTGCGTAGTAAATACTTCATTATAGTGG
TTTGCTTCCAAGAATTTGTTATCTGGAATTGTTAATTGAGTACGTAACATT
AACGCATCGATACCACCACGAACGAACATTAATACGGCAGAAATTAAATAC

20

Sequence 3343
step.1000g08.cons.ok
TGTGCTCTAAAGAAGATCCAAATAACATAGTTATTAGACCAACAAATGTCA
CTGTATAAATATAAACATTACTAAGTCCCAATAAAGGTGTAAATCTAAATA
25 GTAAAAAGATTCCAGCCTTTACCATTGTTGCCGAATGAAGATAAGCACTTA
CTGGTGTAGGTGCTGCCATGGCCTTTGGTAACCAAATATGAAACGGAAATT
GTGCAGATTTGGTAAAAGCACCTAATAATAGCATCAAAATCATTGGTATAA
ATAAAGGATGTCGTGAAATTGCATTGCGTTGATTAAGGATATCAGTAATTG
TGTTTGTTCCTGTAATGATATATAAAATGATAAATCCTGTTAATAACGCTA
30 GCCCACCAAACACTGTAATCATGAAAGATTGAATGGCGCCTAATTGACTTT
CACCATTATTGTACCAATAGGATATAAGCAAGAATGAGGAAATACTTGTGA
GTTCCCAAAATACATACATTAAGATGGTATTATTAGCTATTACAATGCCAA
TCATACTGAACATAAATAATAGTAAATAGATGAAAAATCTAGGAAGATTGT
CCGTACTGTGGGATAAATATTGCGTAGCATAAAAAAATACACCCACACCTA
35 TTAGCGAAATAATTAAGCCGAACATTAAACTTAAACCATCTAATCTTAAAT
CTAAATTAATATCAATTGAAGGCATCCATGGTAATCGAACAGCAATAAACT
TATTTCGAATCACATCTGGTATTTTCATAATAAAATATGTAGATGTCACAA
TAGGAGCTATTAATGCAACATAGCCCGCAACCTTTCTTAGTTGACGATGAC
TCAGAGTGAATAAAACTATGAGCATGACTAATAAATTAGTCGCCATAAGAT
40 ATACTAAACTCATTTAACCCCTCCTTTTCTTTTTACATTTATATTAAGTGT
AATCATTTATTAAGTGTAATCATTGTATGGCGATTTGAAACATATCGTATT
TCAATAAATCTCAATATTTCTTAAGTTATCTGATATCCAAAATATTGTTGA
ATGATAGAAATTAATATTCCCGATTGAAAGAGAAGATAATCAAATGTTTTT
CTTAAATTAGTTAAACTAACAGGTTTTAATCCTTATTATGGTGCAACACGA
45 TTTATGATTTGATATTGCTAGTTGATTTGAATGAGTTTATTGAGCATACAT
AGATTCAAGCACTGTCGCTTTTTTTTCTGAATTGGCATATCACAAAAAGGTA
ACTATTCTTATTTATATCCAAAAGGGAATAGAACTTTATTTTATATAAGAT
AGAATCCATGAGGCTATAATTTTTCAATCATCGATAAACTTAAACTAATTC
ATTCTCATATTGGGAGATCTTAATATTTTAATTCATATATAAATAGTAATT
50 TTAGACACTGATCAATCTTTTAAGTTGGTTATTTTAAATATCAGGTTCCAA
TTTTTTCTTTTACAACTTAACAATTCTACTTACCTCATATACCTTCTAAAA
CAGGTTTCATAATTAATGATAAATGACGTTTGTATCACATCAAAGGCATTT
GCCTATATTTTTCAAAAAATGTGTCCTTTTTTTTTCAATCTATTTCGATATA
TATAGTGAAAGGAGGTTTTATTATGAATGAAGTCAATCAAGTGACGATTAT
55 CTTTACTCAAATCAAACATAATTCTGAAGGTAAAGTAGTCGAACATAAGCG
TCGTTTTTCTAACATTAATCCGCTTGTTTCTAATGATCAAATCAAATCTTT
TCGTTCCATCATTGAACGTATCAGTGGCGAAGAGTATGACAAAGTCGAGAT
CGTTAAAACAGAATCTTTAATTTAAGGAGGATACATCAGTATGTCAAAAAC
TTTGGAACTCGTATTTAAATCAAATCTTAATAAACCCGTCAAATTACTATT

ACCCGATTTTAACGCTATTACTACAGAACAATTAATTAAAGAAAGTATGAA
CCAACTTTTAGAATTAGACATTTTAAGATTTAGTTTGGGAAAGCCTGTCAA
AATTTATGCTGCACAACTTATTGATAAAAGTACAACTGTTATATTTGAAGA
TAAAAATTAATTCATAAAACATTATTTTTGACGCTTTATTTAAAGTACTAA
5   AATCATAAAAAAGACTTATCGAAGAACAAATTTTTTAAAATTGTTCAAGTC
TAAAAAAGAACTTAATTTACATTGTTCTTTTATTAATATGAAGTGAGATAT
CAAATTCAAAATGATGTCTCGCTTTTTTTAATTCTAAAACTTGTTTAATGT
GCATTTTATTATAATGATTCATCAAAATAGTATCGGCTGTATGACACCTCG
GAACATTCACAACTTATTTCTATACTTTGCTTATGGCTCAATCATTTTAGA
10 AACATGATTATGAATGCAGTTTTTCTATGATAATCATAAATTCTATATGAA
ATATCTGATGATTGCATGACTTCATACGCCATAATGCCTGCACCTATGGAA
TGTATCGTTCCAGCTGGTATAAAGTAAAAATCACCTGGTTGCACTTGTATT
GTTCTCAAATAATTTTCAAATGTACCTTGCTCCAATTGCTCTTCAAATTCA
TCACGAGATTTCGCATACGTACCTATAGTAATCTTTGCACCTTCATCAGCT
15 TCAATGATGTACCAACATTCAGATTTTCCATATTGACCTTCTTCGTGTTCA
TATGCATAAGAATCGTCGGGATGTACATGAATAGACAATGGCGCAGCAGCA
TCTACAATTTTAGCCATTAATGGAAAATCTTTACTTGGAAAATCTCCAAAT
ATTTCTCTATGATTGTTCCATACTTGATCCAATGTTTGACCAGCAAATATA
CCATTTTCAATCACACTTTTTCCGTGTGGATGTGCTGAAATTCCCCAACAT
20 TCACCTATGTGATTATTAGGTAGTTGATACCCAAATTGACGAAGATTATCA
CTGCCCCATACTTTATCCAGAAAAATGGGTTTTAAAAATAACGGCATGTTC
ACACCTCCATTATTAAAAGCTTAAACTTAAATCATTATCTATGCTTGAATA
TCACACAATGTTAAATAATAAAACTACTTTTGATACAACATAGCCCTTCTT
TTAAGATATGTATTTGAAGCTAACTATCTTTCTCACTCGTATTGAGTGTAG
25 CATATATGATTATTTCACAAACTACTATATGAATATACTATCATTTTATAA
ACTTTCTATCCTGTGTCCATTTGACGTAGGATTCTTACATATTATATCAAC
TGCCTAATAAATTGAATAGCATCATTTTTACTAATCTAATAAAGCGTTTTA
TTGATTACTATACCATTATTTTCTGAAAATGACCCTTTAAAATAACAATGC
TTTAAATTGGATTCAATAGAATCAAGCATTACACTTTTTTCAAGTTCATTT
30 AAAAGTAGAAAATTTAGTATAATTATACTCATTAAAATATTTAACTTTGAA
TATAAAAAGGGAACAAGATAAGACTTTTGATTTCCTTATCTTGCCCCCTCA
AATAC

35 Sequence 3344
    step.1001a01.cons.ok
    TCTTCGTTTAATTGATAGTTCTCACTAATGAAATCGATGGGTTTTTTATAT
    CTATAAGCTTTAAATAAAACTTTAGCAACTTCTTTATTTGTTAGTGTCATA
    TTGTACCCCTCCCCTAAATTATCTGAAAATTTATACTATTATGATAACTTT
40 TAAATGTGAAGTGTGCAAATTGATTTAAATATACTTCATTTAGGGTATAAA
    ATAATTAATAATGTAAATTTTGATAGGAGTGGGTATAGTGTCAGATTATGT
    TTATGAGTTGATGAAACAACATCATTCAGTTAGAAAATTTAAGAATCAACC
    ACTTGGTTCTGAAACGGTAGAAAAATTAGTAGAGGCGGGACAGAGTGCTTC
    TACATCCAGTTATCTTCAAACTTATTCTATTATTGGTGTTGAAGATCCAAG
45 CATTAAAGCGCGTTTAAAGGAAGTGTCAGGTCAGCCTTATGTTTTAGATAA
    TGGTTATTTATTTGTATTTGTTTTAGATTATTATCGTCATCATTTAGTAGA
    TGAAGTTGCGGCGTCAAATATGGAGACATCATATGGTTCTGCAGAAGGACT
    ATTAGTAGGTACAATAGATGTTGCATTAGTTGCGCAAAACATGGCAGTTGC
    TGCCGAAGATATGGGGTATGGAATTGTTTATTTAGGGTCATTGCGTAATGA
50 TGTTGCGCGAGTGCGTGAAATTTTAAATTTACCTGATTATACGTTTCCGTT
    ATTTGGTATGGCAGTAGGTGAACCTTCTGATGAAGAAAATGGGTCACCTAA
    ACCGCGCTTGCCATTTAAACATATTTTTTCATAAAGACCAGTATGATGCGAA
    TCAGCATCAACAACGTAAAGAATTGGAAGCATACGACCAAGTAGTGAGTGA
    ATATTATAAAGAACGTACTCACGGTGTGCGTACAGAAAATTGGTCACAACA
55 AATAGAAACATTTCTAGGACGTAAAACACGTTTAGATATGTTAGATGAATT
    GAAAAAAGCAGGATTTATTCAAAGATAAAAAGAACCTGAGACAGAAATACA
    TGTCTCAGGCTAGGGTGGGGCGATATTTTCAACACGAATGTTGCCCCGCTC
    TTTTTATAATTGTATGTCGTAATGATCAATGACCACTTTTCTTACATTTAA
    TTTGCAGATAAATCGCCGTAATCATTTGAGTCAAATGTTTTTTTGTCCCAA

```
      TGATTCGTTAAGCGTGCGGTACCTGTTCCAGCTAGCATTGAGTCATTAACG
      TTAAGTGCTGTACGACCCATATCGATGAGAGGTTCGATAGATATCAGTACC
      CCTGCGAGAGCAACTGGTAGATTAAGTGTAGATAATACGAGTATTGATGCG
      AATGTTGCCCCGCCACCTACACCTGCAACGCCAAATGAACTTATAATAACA
   5  ACAGCAATAAGTGTAACAACAAATTGGAAGTCAATTTCTACATTTGCTACT
      GGTGCCACCATAACTGCTAGCATAGCAGGATAGATTCCTGCACAGCCATTT
      TGCCCTATGGATAAACCAAAAGTTGCAGAGAAGTTTGCAATTCCCTCAGGT
      ACACCTAAACGTTTTGTTTGCGTTTGAACATTTAACGGTAATGCACCTGCA
      CTTGAACGTGAAGTAAATGCAAAGATTAGTACTTCTATTGTCTTTTTCACG
  10  TATTTAACGGGATTGATACCTAAGACACTCAGTATAATTAAATGGATAATA
      TACATTGTGATTAGAGCTGCGTATGAAGCAATTAAGAATTTACCTAACGTC
      CAAATTGCAGAAAAATCACTTGTCGCAAGAGTAGAAGCCATAATAGCTAAA
      ATGCCATAAGGCGTTAATCGTAAAACAAAAGTTACGATAGCCATAACGATA
      GAATAGATTGCTTCTATACCACGTTTAAGTAAGCTTCCATGTTCCGGCTGT
  15  TTTCTTGCAACTCTAAGATAAGCAAAGCCCACAAACGTTGCAAAAATAACA
      ACTGCAATTGTCGAAGTTGTACGTTGTCCTGTGAAATCTAAAAATGGATTG
      CTTGGGAATACTTCGAGAATTTGTTGTGGTAAAGTGTTTGCAGTTAAATCT
      TTGGCTTGTTTTGAAATTTCTGTACCACGTGAATGTTCTGCACTACCTAAA
      TCAATAGACGATGCATCTAAACCAAAGATCAAAGCGTAAAAAATTCCAACG
  20  ATAGCTGCAATGGCTACAGTACCAATTAAAAACATAAAAATATAAGAACCG
      ATCTTAGCGAATTTTTCACCAATTTGTATTTTGCTAAAAGCGGCAACAATT
      GAAATGAATATTAGTGGCATGACAATCATTTGTAATAGTGCAACATAACCA
      TCTCCAACAATACTAAACCAGTCTGTTGATTGTTCGAGAGTTTTAGACTCT
      GCACCATATATGAGGTGTAATACAATGCCAAATACGATACCTATTGCTAAA
  25  GCGGTAAAGACACGTTTAGGGAATGAGACGTGTTTTCTTGCCATAATATTT
      AACATTATTAGGAAAACAAGTAACACGATGACATTAATTACTGTAAAAAGT
      ATTTCCATATCGATATATAGCTCCTTTTTAATTTTTAGTATTCCGACATAT
      GTAATAAGAATAAAGGGGGGTTAAGAAGAAATTCAAGTTGTTGATTTATAGT
      TTTCAGATTCATTTAAAGTTGAATAGAGTGAAGGTGATTACTAGAACAATG
  30  AAAAATGAGAAGTTTATGTTTATCAGAAAAGTGAGCAAAGAGGACGTAAAG
      AGGTTTATATTAATAAAAAAGACGCGTGCAATCATGCTATAAGAAGACATG
      ATTGCACGCGTTTTTATTTAATATTATGAATAAACATGAGTCAGCTATGAA
      TACAAATCATTTATAAGTTAAATAGACGTGCAAAGAAACCTTTCTTCTCTT
      CTTTAGGTTGTTGTTGGTAAGTCGATTGAGAAGAAGGTGTAGCTGATGTTG
  35  ATTGAGAATGAGAAATTGACTCAGATTTTTCTGAGTTAGATACACTTCCTG
      ATTCTGACAAGAACTCGCTTGTAGACGTACTTTTTGATTCAATTTCCGATT
      TGCTATTTGTATTACTTTCGGATGTTATTGAGTTAGACTCGGACAGAGATG
      AACTCATTGAATCATTTGCTGATTCAGATTGGCTCAAACTTGCTGTAGTAC
      TTTGACTTGCGACGTTGCTCTGAGATATTGAAGCGCTATTTGACATTGAAT
  40  CAAAATGACTACGTAATTCAAAATATTACATTTAATAAAGATATAACACCT
      AATTACGTTTGTACGTTAATAAATGAAGATTCCAGAAAAATAATTTTAAGT
      AATGTAGACAATCCGAATATCATTATTGAAATGATTCTTATTAATTCAAAA
      AAATTGGTTTTTAATGCTATAAGTAAAGAAGGCTTAGGTACCTCGCCTAAA
      ATCACATTTACAAAATAATTTATACCCTGTAACTCTTTCATTAAAAAAACA
  45  ATATGATATAACTAAAAAGCTCCCAGTAAAGGTCACATTTTAAATGCT


      Sequence 3345
      step.1001a04.cons.ok
  50  ATAAGCTGTTGTGGCAGCAGCAAATGTGCATCCTGCACCATGATTATAACT
      TTGTTGGAACATGTCAGTAGTTAATTGATAAAATTGTTGGCCATCATAGTA
      CAAGTCATAAGATTTATCTTGATCGAGTGCTTTACCACCTTTAATAATGAC
      ATGAGGTGTGCCTTTGTCATAAATCACTTGAGCAGCTTTTTTCATATCCTC
      AATTGATGTTAATTTTCCTAAACCAGAGAGTTGACCTGCTTCGAATAAATT
  55  CGGGGTAACAACTGTAGCTTTAGGTAGTAAATATTGAATCATTGCTTCTGT
      GTTTCCTGGGTTAAGTACTTCGTCTTCTCCTTTACAAACCATTACTGGATC
      AACTACAAAATAGTCTGCACCAGATTCAACAAAAACATCTCCGGCACGTTT
      AATAATGTCTTGTGTCCCTAACATTCCTGTTTTAATAGCATCAGGTCCAAT
      TGATATTGCAGTTTCAAGTTGTTTTTTCGAAAACATTCATATCAATAGGTGT
```

```
      TACATCATGGGACCATGTTTCTTTATCCATTGTTACAATAGATGTTAAAGC
      GACCATTCCATATACATCAAGTTCTTGGAACGTTTTAAGATCTGCTTGCAT
      ACCGGCACCAGCACTTGTATCTGAACCAGCTATCGTTAATACTTTTTTTAA
      AGCCATCATTCATTCACTCCCATTAATTTCTAGTGTCTTTATCATATCATG
  5   TTTATCGCGTACGCTAAATTATTATAATTTTAAAATGCAAATCAATCATCA
      TACTTAGCTATTGTATTGATGTTATCTTTAATTCGTGATTAATAGGTAAGA
      ATGTCATCTCATTCTTACTTAAAAGGTCTTCAGTTAATTTTATTTTGAAAA
      TTTGATTGCTAAAGTAAGGCTTAATATAATAACTTCTTTCTTTACTTTCCA
      TCACTAATTGATATTGTGTGTAATGTATTTTATTGGCATCGATAACTGCAC
  10  CTTGAGGGATACTGACTGATTCTAGAACTTTAAAACAATTCATTAAATTTT
      CATCTTCATTATGGGAACAGCGTAGTTGGTGTCTTAAATATGTAGCGCGTA
      TAAAACGATCTGTTGACGTATAACCACCCGGTAAGCCATTTGTTCCTGCTT
      CACAGCCCATTGATCTTACTAGCACTTTACCTATTAATTGATTGGTTGATT
      TCTGTGGCGTTAAAAAAGCGTAATTTCTTAAGTTAGATAGATGCCAATCTA
  15  ATTTAGGTTCATTTGTTAAGGTATGAACATAATTATCTTTAACTATTAATA
      AGCCATTGTGAGGTTCTATGGCTACGGTATGTCCTGTTCATCAGTGACCA
      TGAAATGTAAAGGAGGAACGATATTCAAAGTCGTATTTTTTTCATTCATAA
      TATTGATTTTCTTAACCTTTTGTTTTAATTCGCTAATACTTTTATTAAAAC
      CTAAAACCCAAACAATAAACTCCTCAGGTGCTAAGTTAAAATAACCATAAC
  20  GTTTATGGGTACTGTATGAGGCTTCACCAGTGAAGTAATGGTTCGAAATAG
      CTAAACCTTTTTCGTTTATACCATCACCAAATCTATAACGTCCTACTTTTA
      AATTTGTTCCAACAAACCATATTCAAGACGCATGTCTGAATCTAGATCAA
      ATTGGTAGTGATAATGGCGTGGAACAATGGTTGGGATACCATTAAATTCAA
      ATGCAAAGTCCATTGTTCTAGCTAAATAATGGTAACGTTGTTTTGTATATA
  25  AAGAAATGGCAGTACACATATATAACACTCCAAACTACTGAAATTTATCTA
      TATTAATTGTAATATAGATAAAAATGCGCTTATTGTAAACCTGAATCATAT
      AAAAGTTTACAAAAATTAATGCTTATAAAATCAGTCATTTTTCTAATATGA
      ATATGTTAAAATGTAGCTTGAGGTGAAAACGAATGAAATGGTCAGAGGTAT
      TTCATGATATAACAACGCGCCATGATTTTCAGGCGATGCATGACTTTTTAG
  30  AAAAAGAATATACGACTCAAACCGTCTATCCAGATATACAAAATATCTATC
      AAGCATTTGATTTAACGCCGTTTGAAGATATCAAGGTTGTTATTTTAGGGC
      AAGATCCTTATCACGGTCCTAATCAAGCACATGGTTTAGCATTTTCAGTGC
      AACCTCATGCTAAATTTCCACCATCTTTAAGAAATATGTATCAAGAACTAG
      AAAATGATATAGGGTGTCATAGAACTTCGCCTCATTTACAAGACTGGGCAA
  35  GAGAAGGTGTCTTGTTATTAAATACGGTATTGACTGTTCGACAAGGTGAAG
      CACATTCACATCGAAATATTGGATGGGAAACATTCACGGATGAAATCATAC
      AAGCTGTTTCTAATTATCGTGAGCATGTTGTTTTTATTCTGTGGGGAAGAC
      CGGCTCAACAAAAGGAACGATTCATTGATACATCTAAACACTTAATCATTA
      AATCGCCACATCCTAGTCCACTATCGGCTTTTAGAGGATTTTTTGGTTCTA
  40  AACCTTATTCAACTACAAATAACTATTTAAAATCTAAAGGGAAAACACCAG
      TTCAGTGGTGTGAAAGTTAGGTGAGAATAATGAACAAAGAACAGATTCTAC
      AATTGATTGAGCAAGAATTGATACAAGCAGATGAAGCTCAGACAGATACGG.
      AATTTGAAAAGCATATGTATGCTATACACATGCTCACATCTCTTGTTAGTT
      CTCATCAAAGTCGTTCTACAATAGAGAAATTAAATCATTCTAAACCAATGA
  45  ATAGTAATATCAAAGATGATTATGAGATGAAACAACAGTCTTCACAAAAAC
      ATCATGTAACTGCAGCTGAAATAGAAGCAATGGGTGGTAAAGTACCACAAT
      CAATGAAAAAGCATCATACTTCTAATAATATGATGATTACAGATGATCAAG
      TTGGTAATGGTGAATCTATTTTTTGATTTTTAAATATAAAGGAGTACATTCA
      AAATGATGAAAGTTTTT
  50
```

```
      Sequence 3346
      step.1001a09.cons.ok
      CTTCTTAAATTACGTAATTTGTATCCAATATCCATCTTTCTAACCACCTAA
  55  AAAAGCTATATCATGTTTACTTATATCAAACATTTTGTTTAATGCTTAATA
      AAAATAACAGAATATAGCTACGTAATCAATAGTTTTATATTAACTTATTAT
      TGTGGATTTAATAGAAAAAGTTAGTCGAAATCTTTACTAAATAAAGTTTGA
      TTTGGATTTTTTTTAATTTCTTGAACTACTTTGTAACTATTATAACCACTT
      TTATTTTCAAATTCTTTAAAAAATCTGTTTTTCATCAGCTTTTCCAGGCACA
```

```
ATATGTTTAAAATTTTTATAAGCATTTAAAACTTCGTCACCTTTTACAGAT
GCCTCATAATAATTTTCGATTTTATTAAAGAAATGAATTACTTTCATCATT
TCATCATTAGTCCAATCTAAATCTAATGGATATTGATATTCCATAGTGTCC
TCTCCTTCATATACTTTTATTTTTCACATTTAGATATATTAAAAATAACAC
AGTATACCTTTATTGGCATAAAAAGATGGGATAAAATCCATTTTTGAATTC
TATCCCACCATGAATAAATACCCCTAGCAGATTGCTATTATTTAGAGTAGC
TAAACGCTAAGGGTATTTTGAAATGATTTAAGATACGTATGTCTTTGCCAG
TCCGTGACTATACAAACGAAAAGTAATATACGTACAAATCATAAATAAAT
CATATTAATTTATATACATTTATCATTATTAACTAGTAATTACATAGTATG
AATTGGATAACCAATTGCTTTTTCAGCAGCTTCCATTGTCATTTCACCTAA
AGTTGGGTGTGCATGTACAGTTAATGCGATATCTTCAGCATTCATACCTGA
CTCAATAGCTAAACCTAATTCAGAGATAATATCAGATGCGCCAGTACCTAC
AACTTGTGCTCCAATAAGCGTATCATCTTCTTTAAGTGTAATTAACTTAAC
AAAACCATTTGTATCATCTAATGATAAAGCTCGTCCATTAGCTGCATAAGG
GAATTTAGAAGCTTTAATTGATAAACCTTCTTCTTTTGCTTGAGCTTCAGT
ATAACCAACTTGTGCTAATTCTGGTTCTGTAAAGCAAACTGCTGGCATACC
AATATAGTCTACCTCTGCGGCTTGACCATCTATCGCTTCAGCAGCAACTTT
ACCTTCATAACTAGCTTTGTGAGCTAATGGTAATCCAGGTACAATATCTCC
AATCGCAAAGATATTTTCAATAGAAGTACGACTTTGTTTGTCCACTTCTAG
TAATCCACGATCAGCAAATTTCAGACCAAGTTCTTCTAATCCTAATTCATC
AGTATTAGGGCGACGGCCAACTGTAACTAATACATAATCAGCTTCGATAGT
TTGTTCCTCACCTTTTGCCTCATAAGTTACTTTGACACCATTTTCAGTTTC
TTCTGCAGATTTTGCCATTGCTTCAGTAACGATTTCGATACCTTTTTCTTT
CATACCTTTTTTAACAGGTTGTGTCATTTGCTTTTCAAATCCGCCTAAAAT
ATCTTTTGCACCTTCAAGGATAGTAACTTCAGAGCCAAAGTTTGCAAAAGC
AGTACCTAATTCAGAACCGATATATCCGCCACCAACTACAACTAGTTTGTT
AGGTACTTCTTGTAGATTTAAAGCTCCTGTTGAATCGATAACACGTTTACC
AAATTCAAAATTTGGAATTTCAATTGGTCTTGAACCTGTAGCTATAATCGC
ATGTTTGAAATTGTAAGTTTGAGCACTCTTTTCGTCCATGACACGTAAACT
ATTGTTATCAACGAAATAAGCTTCACCTCTAACAATCTCTACTTTGTTACC
TTTTAAAAGTCCTTCAACACCGCCAGTTAATTTATTAACTACAGAAGTCTT
GAATTCTTGAACTTTTTGATAGTTTAACGAAACGCTTTCAGCAATTACCCC
TAAGTTTTCTGAATTTTGCGCTTCAACAAAGCGATGAGAAGCATGTAGTAA
TGCTTTTGAAGGTATACAACCAACGTTTAAGCATACACCACCTAAATTACC
TTTCTCAACGATTGTTACCTTTTGTCCTAATTGAGCCGCGCGAATGGCTGC
GACATATCCACCTGGACCTGCTCCTATTACAATAGTATCTGTTTCAATTGG
GAAATCTCCAACTACCATGTTTTTACCCCTCCATTAATAATAATTCTGGAT
TATTTAATAAGCGTTTAATGTGATTCATAGCATTTTGTCCAGTAGCACCAT
CGATTTGTCTATGGTCAAAGCTTAATGATAAAGCTAACACTGGTGCAGCTA
CAATTTCTCCATCTTTAACGATAGGTTTTTGAGCGATACGGCCAATTCCTA
AGATAGCTACTTCTGGGTGATTGATAACTGGAGTGAACCATTGTCCACCAG
CGGAACCGATATTACTAATTGTGCATGTTGCACCTTTCATTTCTTCTGAAG
TTAATTTACCATCACGTGCTTTTACAGCTAGTTCATTAATTTCATCAGAAA
TTTCGAATATTGATTTACGATCGGCATGTTTAACTACTGGTACTAATAATC
CTTTATCCGTATCTGCAGCAATACCAATATTCCAGTAATGTTTGTGTACAA
CCTCTCCAGCTTCTTCATTGAAAGAAGTATTAAGTGCTGGATATTTTTTAA
GTGCAGAAACTAATGCTTTAACAACATATGGTAAGAAAGTAAGTTTTGTAC
CTTGTTCAGCAGCAATTTCTTTAAATTTCTTACGGTGATCCCATAATTCTT
GCACATCAATTTCATCCATTAATGTAACATGAGGTGCAGTGTGTTAGAAT
TAACCATTGCTTTAGCAATTGCTTTGCGCATTGCAGGTATTTTTTCTGTAG
TTTCAGGGAAGTCGCCTTCTGGTAATGCTTGTGTTGCAGAAGCATTAACGA
CATCACTAGAAGTTGATTCAGATGCTGCGCTAGTGTTTGAACCTTCTTCGG
AACTACCACCATTTAAGTATGCATCGATGTCTTCTTTTGTGATTCGTCCAT
TTTTACCAGAACCATTTACAGCTTTAATATTGACACCATTTTCACGTGCAT
ACTTACGCACTGACGGCATCGCTTTAACAGTTTTACTTTCATCTACTTCTG
TCTTTTCTTGTGATTGAGTTGATGAAGCTTCTTCTTGCACTGGTGATTCTT
GTTCTTTTTCTTCTTTCTTAGAATCCTCATCATCGCCATGACCTTTAAATT
GCATTTCTTCTGCATCAGGTGCATCAATTTTAACGATGACATCTCCTACTA
CTGCCACTGTTCCTTCATCTACTAACACTTCTTCAACAGTACCACTTACTG
```

GAGAAGGAATTTCTACTACAGATTTATCATTTTGAACTTCTGCTAATACAT
CATCTTCTTCAATTGTATCGCCGGCTTTAATAAACCATTTAACAATTTCAC
CTTCGTGGATACCTTCCCCGATATCGGGTAATCTAAATTCAAATGCCACGT
TCTTGTCCTCCTAAGATTTACGTATTTAAATTT

5

Sequence 3347
step.1001b03.cons.ok
AATTTTAACCTCTCAACAAAAAGTCGTGATGGAACAATTTTTTACACCTAC
10   ACAAAAAGTTAAATTTCAAAAATATTACAACCCTGAGCACGAACATCCTAC
GGTACAATCTATCATTTATAATACTTCACGAGACGTTCGTTTTTTCAACGA
TGAAAATGAACTTTTAGCGTTTGCAATTAATGCGCTATATCATTTAGGAGA
CGTATTTTTATGTGATAAAAACATCGTTACAGGGCCAATCATTGATCAAAC
TGATACTAAAATACCAGTTCTCGCCGTTTTCCACAGTACCCACGTAAAAAA
15   TATCAATGATATATATCATTCTGAAATCAAACAAGCGTATAAACCTGTTTT
AGATAACTTATCCCGATATTCAGGAATCATAGTATCTACTGAACAACAAAA
AACAGATTTATCTGTAAAAATTAATAACGTTATTCCCATTTACGTTATACC
TGCAGGTTATATTGATACAAATGAATCTCATCATAGTAGTGACAATAAACC
ATTGCCTAACAAAATGATTTCTATCGGCCGTTATTCTCCTGAAAAGCAATT
20   AGATCATCAAATAGAACTAATGTCTAAGCTAGTTCCAGCATTTCCAAATTT
ACAGTTACATTTATTTGGGTTTGGTAAAGAAGAAACACATTATCGTAAATT
AATAGCTCAATATCATTTAGAAAATCACGTGTTTTTACGTGGCTTCATTTA
TGATTTAAATCAAGAAATAGAGACCGCCTATTTATCTTTATTGACAAGTAA
AATGGAAGGGTTCAATTTAGGTGTACTTGAAACGATTGCTAAAGGCGTACC
25   TACAGTAAGTTATGATACCAAATACGGACCTTCTGAGTTAATTGTTAATCA
TAAAAATGGCTTTTTAATTGAACAAGATAATAAAGAACAACTCTATCACAG
CGTTAAAAAGTTATTACTCGATTCTAACTTAAGAGAACAATTTTCTAAGGA
AAGTATTAAACATGCCCAAATATTTAATGACAAAAATGTTTTTGATACATG
GCTCACTGTTTTCAGAACGTTAAAAGTTAATTTATAATCGCCAATTTATTA
30   AGATATATAAGGGTGAAGAGAAGTGTGAAACATCATACATCATCTTCACCT
TTAATTTATTCTATAATATTTCATCTTTAATCTGTGCCTTTTGATATTGAA
CATAGCCTGATATATAATGTTCGATATCATCAATCCGAACTCGATACCCAT
GGTGTTTTATGAAGAAACTACCAAGAATTCGCTTTGGATTTTTGAAGTACA
TTGCAATTTCAGGATAGAAAAATCCTACTCGTTGATATTCTGCTCTTGTAT
35   GAATCACATTGATTAATTTTTCTTCATCAATCAACTGAGTCACAAGCTCTT
GACGATGTGTAGCTTTTGCTTTTTGAATAAGCTTATAGGTTGCCATTAACA
TTTCTAAAAATGTTGGGAATGTTGTTTCACGTTGTTCAATATATTCCAAAT
AGGTGTTCACATTCTTAATCCCAAATTCAAAATATTTATCTTGCGGACACA
ATTGAACTAACTCGTTCGTGCAATACCCAAGCCAATGATCATGGTATTGCC
40   AATACTCTTTTTCAATAAATCGATCCATTAACTTCTTCACAACTTCCAACC
ATTTATCATTATGATCTTGGTGATATAAGCGTAATAATGCTAAAGCTGCTT
CACCATCATAATAAATAATTCTAAATGATTCTTTCACAGTTAAATCCGGAT
AATTTAAAATATGAGTTGTTTCGTATGTATCTTGATTAATCATTGATAAAA
TTCCTCGAGCAACTTTTTGTGCCACGCATAAGTATTGCTTATTGGGGTTAT
45   GCTTTAAATATTCACAAACCGCAAATATAAAGGCAGCATTTTGTCCTAATT
TTATTTCGTTAATATCTTTTGTATCATCAAAGATATATCCAACACCTTCAT
TATCATAGAAATAATTCTCAATAACGTAGTTAATTGCCTTTTCGACTATAG
TTAAATCTTCTCCTAAATAATCTAAACCCTCTATTAATGCATAAGTTGAAG
AAGCATGTCTTAATATATTATAGAAATTGATTTCTTTATCAAAATGTGGAA
50   AATAACCATATTGATATCTTCCAGTATCTGATAGCATATTTCCTAGGAAAT
ATGTACCACTTTCAATTAATTGGTCTATTTCTTTATGTAAATAATCGACCT
TTCTTAATCCTTTTTTATAACCTTCATCGTGTAATTCATATATCTTTTGTT
CGTCTAAGATAAAACCTTTAGTTTTGAACTTAATGACTTCTTTGTTTTCAT
AAAAATCATAAGCAAACTTTTTCTTATGGTTCGTATACTTTCTTAAATAGT
55   TATTTATATTTTGTTCAGATAGAATAAGCTTCGTTTTCCCATCTGTTTTCA
CTGGTTTAATAAATGCATTAGTGTTTATTTCTTCAGGTAAAAATGATAAAT
TCCAGTATTGATCTAAAGCTATACCAAAATCAATATAATTTCTTCTAGTTT
GCGTTAATTCATCTTTAACATCTTTAAATAAAGTAACTTCTTCTTCAGTAA
CAATGTCTATTTTTACCCATGAAGGAAACGCTCCTGTTTTCTGTCTAAACT

```
        TTTGAATTTCTTTAGTAATATCCCTTCTAAGGTAATTAGGCTTTTTTAATA
        ATTTAACTTTGGCTTTTATATTAGATTTTCCAAGGCTTAAAAAAGCATATT
        CTTTTATTTCTATTGTTTTTTTCAACATGATTAATCAATCTTTCTATATGAT
        TTTCCATACTCATCTTTTTCTCCTTTCAATTTAATATGATTTCAAACCAGG
    5   CCATATAGACCTGGCTTGAAATAACTTATGAATGAATCAATAGCGCTATTT
        TTCTTTATTATCTTTTTTACGACGTCTTCCTAATAATAATGCTCCTAAACC
        TGCAAATAAAGTTCCAAGTAATGTGCCTTTAGAATCATGATCTTCATTTGC
        TCC


   10


        Sequence 3348
        step.1001b04.cons.ok
        CGTGACTATATGATGGTAGAAATGTGTCCTCGTTGTAAAGCGCATTTTAAC
        AGTCGTTGTAAATTTCATTATCATTTATATTTTGAAGTATAAATAACTTGT
   15   TACAAAGTTTTTGAAATATTTGAGCTCCTAAAGTGAATTTTATTTAAAAAT
        TCGCTTTAGGAGCTCTTCTTATATCAACATTCTATTCATGCATTTAAAGTT
        TTCTTCTTATACCTCGGTTGATTTTCTCTTATTATAAGATTATTTTTCAAC
        TTTCTCTACTACTTCATTTAATTCATCAATTTGTTTGATAGTTGTCGTTGA
        AGATCCTGAAGAGAAATACCATAGTTTTGGATCTAATTCGTAAATATGATT
   20   ACTTTTTACTGCTTTTACATTTTTTATAACTTTGTTTTTTAAAACTTGATT
        TGTTGTTGCTTTACCACCTACAACTGAACCACGATCCATAGCTAAAATAAC
        ATCTGGATTCTGCTTGTTAATATATTCATTATTTATATTTTGACCATGCGG
        GCTTTTGCTAACCTTTTTGTCTGCAGGTTTAAATCCTAATGTATCAAACAC
        TAAACCACCAAATCTTCCTCCTGGTCCAAACGTTGATAGTTCACCTTCGTT
   25   AACCAATAAATACATTACTTTCTTATTAAAGTCTTTAGTTTTATCTTTCAT
        ATCAGATATTTTTCTATCTAAATCTTTATTAATTTTTTTAGCTTTATCTTC
        TTTATCGTAGATTTTCCCTAAATTTTCTGTATTTTTTTTCATATCTTTAAT
        TAAGTTGTCATCACTTGTACCTACATATACAACTTTAGCTTTTGGTGCAGC
        TTTTTTAAATTCATCTAAATTTTTCTGATTAGCTGTTCTTCCTGAAATAAA
   30   AATCACATCTGGTTTAGCTGATGCAACTTTATCAAAGTTCACTTCTTTTAA
        ATTTCCAGTATTAATATACTTATCATCTTTAAATTCATCTAAAAATTTAGG
        TAAAGATTGGTTATTTTCACCTTTAGGTAAACCTTTTACTTTATCAGCCAC
        ACCTAATTCTTTCAACACATCAAGCGCTCCATAATCTAATACAACGGCATT
        TTTAGGATTCTTTGGTACTTCGACAGTATTAGAGATTTTTTTCTTATCACT
   35   GCCATTATTTTCTTTACCACTTGCTTCAAAACTATTTTTGATGGTTACAGT
        TTCTTTAGAATCACTATTTTTCTTTTTCGAAGTTGAATTATTATTCGAACT
        ATTACTACAAGCCGTTAAAACTAAAACTAGAGACAATAATAAAAATAAGAC
        TGTTTTTTTCATATGTGCAACTCCCTAATCAAATTTATAATTTCTTTTCGT
        ATGGATAACCTCTTTTACTATGCGTCTCAATTTTAATCATTAAACTGAGTC
   40   AAAAGTAGTTTCATCATAATATAGACAAATACGTTGTCCCCTTATCTCCTC
        TATACGTACTTCCATTTCATATAAACTTTTTAAAATGTCAGACTGAATAAC
        ATTATCTTTATCATCAGCTTTAACTAGCTCACCTTGTTTAAGCGCAATGAT
        GTCGTCTGAATAACAAGAAGCAAAGTTGATATCATGTAATACGATAATGAT
        TGTTTTATTTAACTGACGACATAAATCTCGTAACGTTTGCATAATCTGTAC
   45   TGAATGTTTCATATCTAAATTATTTAAAGGCTCGTCTAATAATATATAATC
        AGTATCTTGTGCTATAGTCATAGCAATATATGCACGTTGTCGCTGACCTCC
        AGATAGTGTTTTAATATTACGATGTTTAATTTCATTAAGCTGTAATAAGTC
        AAGAGCATATTCCACTTTGTCATAGTCTTCTTGCTTCATTCGACCTTTTGA
        ATATGGAAAACGACCAAATTTAACTAGTTGCTCGATAGTTATGTTCATATC
   50   AGTGTGATTAGATTGTTTTAAAATACTTAATTTTTTAGCTAAATCGTCAGT
        TTTATATTCTGAAATTTCTTTGTCTTCAATCAAAATTGAACCGTTCTCATA
        ACTAAATAAGCGACTGATAGCAGATAATAATGTACTCTTACCAGCACCGTT
        AGGACCTATGAGTGAAGTCAATTTTCCCTTTTTAATATCTACGCTAATATC
        TTTTAAAATAGGTTTATTGTCTATCGTTTTATCTAAATTTCGAATTTGAAT
   55   CAATTTGCATTTCTCCTTTTGACAAGCAAATATATGAAGTAACTTCCACCA
        ACTAAATCTACTATCAAACTGAATTCAGTTGTTGCTTCGAATAGATTTTCA
        ACTACCCATTGCGCTATAAATAAACTAATCCAGCTGAATAATATTGTTGCC
        GGTAAAATAAATTTGTGTTCATAGGTCTTCATAAATTCGTGTGCTAAGTTG
        ACAGTTAATAAGCCAAGAAATGTTACAGGACCTATTAATGCAGTAGATATA
```

GAGACGAGTAAAGCAACCAAAATAAGAAACATACGCGTCATATTCTCATAA
GAAACACCTAGATTAATTGCTTGTGCTCTACCTAACAAAAGAACATCTAAA
TATGGTCGAAGTATTATTGTAACAATTATTAATATAACTAATAATATGCCT
GAAACAGTTACTAATTTAGAATTTGATGCTTCAAAGCTAGCAAACATAACA
5   TTTTGAACTGCTAAAAAAGATTCAGGATTCATTATAAGTTGGAGAAAACTT
GTAATACTACGGAAGAAAGTTCCTAATATTACACCAACTAGTAATATGAAA
TAGACTGAAAAGTGACCTAGTTTAAATATAACTTGAAATAATAGCAATGAA
AAAAACACCATTGCGATCAAAGTGATTAAAAAATTAAGGTAGATATTTGTA
ACAACAGTTGCTTGTTCACCTAAAATAAAAATAGGTAAGACTTTGACAAAT
10   AAATACACAGAATCAAGTCCCATAATAGAAGGAGTTAATAATCTATTTGTA
GTAATAGATTGGAAAATAACTACGGACGTTCCTATCGATGCTCCTACTAAT
AATATAAGAATAAATTTTTTGAAACGGCTTTGGAATTGATATTCAAATATA
TCAAAATCTAAACCTACCAATAAATAAAAAATTGCCATACAAATTGCCATA
ACGACCAGTATACTTATTTTTTTATTAGCGCTCATTTGCATAATTTTTCCT
15   ACCTTTCATTAGCAAGATTAGGAAAATAATTGTGCCAAATAC

Sequence 3349
step.1001b10.cons.ok
20   CAATGGACTTTTATCGACTTGAGTTTCAGAAGCAAATATAAAAAAAGAACC
TCGAGCCATACCTATTGTTCCAATAAATAATAAAGAAATAACATTAATATA
TTCCATTTCAGTCATACGGTGTATTTTCTTCATGTTTCTACCCACTTTCTA
TCATTAAGTGCAGCACAAATAGCATTAGCATAATAATACATATCTTCTGTC
CAATTATGTGTATTTTCTCTATTCTTTTTACTAAAATCGTTTCGACTAAAA
25   AATTCTTTTCTTTGTTGAGTCGATATTTCAATCTGTACACCCATGCCATAG
GCATTTTTGTTGGTTATATTATTAATTTCTCTGCCCGCAATTCTGTCAGGA
GCAGCTTCCACATTAAAACCTGAAACTTTTAAATTGTGAGTAATAAGAGAT
ATAAGTCTTTCATCCAATCCACCAATATAACTATTTGCTTGATTGCTCGAA
TAACCATGTACGGCTATCGTTACATTCATAAATTGATTCCAATATAATAAA
30   TTGGGGTTATCATAATTTGTTGAAGTGACGTGTAGAGTTCTATTGTTTTTC
GGTTTTAAACCTTTAAAAGTAAAATAGTTTGCATTCGATAATTCTGCAACC
AATAACGCTAATTCAGAAGTGCCACACTCTATACCGCCTCCGTGTATAGCA
GTTATAAGTGATTTACTATTTCTATCTTGTGTCTCAATCATCCAATCTTTT
TCATTCCTCACAAGTTCAGTCATAGATTTATAAGTATCCATCACTTTATTC
35   ACCTCCTTTATCTATCATTCTAATTTCTCCGTAAATATAAGCTTCACGACT
TACTGACCACTGATCGCTCTGAGAAATATAAACTTTAACATCTCCATTAGG
TTCAATAGTTAATTGAGCACCACAAGCTTTAACAGGCACCGATCTTAGGAA
CGCAGTTTGCGTACTTGTAATTAGTTCACTCGGTAACTTCGCTATAACAGT
TCCACTTTTAAAGTTGTCAGCGTTAATTCTTAACATCACTTCCTTATAGTC
40   TTGATGTTGAATGATTTTATAAGCGCAATTAAAACCATTTTGTCCTTCAGC
TTTATAATGCCTATTTTTTATAGCTCCATTAATCAAATCATATTCAATCCA
GTCTGAAGATTCAGGTAACTTTTCTTGTATTTCAGTAATGCTTTTACGATG
TTCATCAATTGTGACGTATCCCTCTTGACTTAATATATTTTCAATATTCCT
AACTTCGATTCTTAAATCTTTTAAATATTGACTTCCTTTTGTGTCAATTTC
45   GTTGTCATATTCTTCATAAAGATTTTGAATTGTATTTTTAGCCGCGTTAAG
AGATTCCTCGACGTCTTTTAAACTTTTACTACGCTGTTGATTAAGTTTGTC
CAATCCTTTTTTTGTTTCAATCTGAATCTTTGTAATTCCTTCATCACTTGC
ATCTTTCACTTTAACAACGTAATCTTCTAAATTATCTAATTGTTCTTGTAT
TTCAGTGGCTCGTGCATTAATTTGTCTTTTTAATTCATCAAACATACGGAT
50   ATATTTAATTTTAGTTGCACCATTAATTTTATTTATTATTGCGTCACCCAC
TTCGAATTCAAATTCAGTTAAAACTGCAGTTGATGATTTATCGCTATCTAC
CTGATTTCGATGGTTAATCGAAATATATATTTGACCTATAACAGTTGTATT
CGTTGAAGCTTGTAAGAATTCAATAGGTATAGTTACTTCAATAACACCATT
TAAAGGGTCAATAAAACGTACATTTTCTACTACTTGATTAGAACCGTTAGA
55   GGACTCAAGATATATATACGTTTCAGTGTTTTCTTCACTAATTAATAAAGG
CTTTTTGTTTTTAGTCACATAAAAACCTTAAAACTGCAGTTTTATCATCTAA
GTTATAAAAACCGATACCTTCATCAGATATCGGTTTTAAGTAAGGTTCATT
TTTAACTTCAATTTTACCAATCTTTTTTTAATTCCATTATTTTTCCTCCCCA
TTTATAACATTTCTTAAATCATCATCATATATAGTATTAGGGTAAATTTGT

```
GTGAACGTATCTTCTTTAAGATTACCGTATTCTCCAGATTTTAATATCTGA
ATGGAATTAGCAGAATGCGTAGGAGTGAACTTCGCAAATAATTTAATTTGT
TTGATTGTTACAATACCAGCACCTTTTTTCTTAACATCTAGTACCGGCATT
ACTGTATCAGTACGTTTTGTACCTGGTTCAGTAATCGTCGACATTTGAACT
```
5   ```
CCAGCAGCAGCGTAAATAGGGAAAGTTGTATTTCCTTTGTTGAGTCTATGT
TCAATTGAAAATAAATTACGTTTTCTACTTTTATTAAATCCAAAGAATGGA
TGATAATTCTGAACAATATTAGGATTAATCCCCACTGTGACATCTCTATCA
ACGTTTATAGTCACATATCCATAGAGCTCTACAAATCCATTTGCTAAAACT
TTGAAACGCTGCTGACTCATCATTATACGTTGCCATTCATTCCTTTCTGCA
```
10   ```
CGTAGGGTAATGACTTCAGATGTTTATTGTTAAATCTATCATCATATACC
ATCACTTTAATGAAAGGATCTAATGATGTTTGACCATTTTGATAAGATTCA
AATTGTGCCATTCTAAAAATTACATTACCTACCCAGCGTACTGAACGTCTC
ATTTCTTCGGGACTTTTCTTTTCACCGACACGTTTTTCATTTATTTCTGAA
AGATATGACGTTGTTTTATTTCGGATTCCTACCCAATTACTAAAGATGAT
```
15   ```
AATGTACTGGAACCCCAAACAACGAGGTCTCCGTTACTTTTTAAATTGTTA
ATAAGATGAGTCATATTATTATTTTTTTGATTGGCCCAACGCGGATAAAAT
AAACAGTAATCATTATTTACTGAAATAATATCATGCAAATCTAAGTGAGCA
GTTAAATTATCTATACTTTGAACTAAGCTACGCATATTTTGTG
```

20

```
Sequence 3350
step.1001c01.cons.ok
CAAAGTGAGAGGCAACACAGGTACAATGGCTCTTAAATAAAGCATAATGAA
CCCAGGAATATATCCGAATTGTCGAAATGCGTCAAACCATTCTTCTAACTG
```
25   ```
AGAAAGCGACAATACATCAAATCCTTTCTTAGGGTTGTATTACTTATTCAT
TATAGTTGTAAATGATATTTTTTGAACATTAAAATTAGGTCAACAACATA
AATCAGACTTAGGTAGGAATTAATAAATCTATATATTGATATCTAATTCAT
CAATTTTATTATTCCATTCTTTATCGAAACTCTAATGAATTGAAAATTTAA
TATTTCCTATCTGAATTTGACTTCTTATTAATCTGACGTTAATCAATTTTC
```
30   ```
ATTTACTCTTATTAATACACATAATAAAAATGAGCCTAGATAATCCTAGGC
TCATTTTTAAGTTCAACTTAAAATTATAAACGATTTTCTAATTCTTTTTTC
TTATCTTCAAAGCCTGGTTTACCTAATAGCGCAAACATATTTTGTTTATAG
GCTTCAACTCCAGGTTGATTAAATGGATTAACACCTAATTGATATCCACTC
ATTGCACAAGCTAATTCAAAGAAATAAACAACATATCCAAATGTTTCTTCA
```
35   ```
TCTAACTGAGGAATATTTACAACGATATTTGGAACGCCACCATCGGTATGT
GCAAGTAATGTACCTTCAAATGCTTTAGTATTCACTTCATCGATTGATTTG
CCAGCAAGATAGTTCAGTCCGTCTAAATCATCTGCATCCTCTTCAATTTTG
ATATCATGTTTTGGATGGTTGACCTTAACCACTGTCTCGAATAAGAAACGA
CGGCCCTCTTGAACATATTGTCCTAAGGAATGTAAATCAGTTGTGTAATTC
```
40   ```
GCACTTGATGGATAAATACCTTTGAAATCTTTCCCTTCTGATTCACCGTAT
AATTGTTTCCACCATTCGTTGAAATACTGCATAGAGGGTTCGTAATTAATT
AACATTTCAGTAGTATAACCTTTGCTGTATAAAATATTTCGAATAGTTGCA
TATTGATATGCGATATTTTGATCTAAATCATCAGAAGATAACTCTTCACGT
GCCTTAGCCGCACCAATCATGATTGATTCAATATTGATACCTGCAGTTGCA
```
45   ```
ATTGGTAGTAATCCTACAGCTGTAAGAACAGAATAACGACCTCCCACATCA
TCAGGTACAACAAACGTCTCATAACCCTCATTGTCTGCTAATTGTTTAAGT
GCACCTTTAGATTTATCTGTCGTTGCAAAAATACGTTTCTTAGCTTCATCT
TTTCCATATTTTTCTTCAACCAATTGTTTAAATAATCTAAATGCAACTGCT
GGTTCTGTCGTAGTACCTGATTTTGAAATAACGTTAACTGAAAAATCTTTT
```
50   ```
CCTTGTAAATAATCAAGTAATTCTTTTGTATAACTTGAGGATAAATGATTA
CCTACAAATACAATTTCAGGGTATTCCGTATTTGTTCTAAATGAAGATGTA
AGCATCTCGATTGCAGCACGTGCACCTAAGTATGAACCTCCAATACCGATA
ACAACAAGTACATCGGAATTTGATTTGATACGTTTAGATGCTTCGACGATT
CTAGAAAATTCTTCTTTATCATAATCAACAGGTAAATCTAACCAACCTAAA
```
55   ```
AAGTCATTACCTGCTCCTGTACCTTTATGAATAGTTTGATGGATAGTTTTA
ACAATATCCTTTTGCTGATCTAGTTCATGCTTATCAAAAAATTCTAAAGTT
TTGCCATAGTCTAATTGAATGTGAGTCATCTTATTGCCTCCTGTTATAAGT
ACTTATACTAATTCTACTAAGTTTAGATGCGCTATTCAATTCTCAAGGCTA
ACTTATCACAGGAAAATTATTTTTTATGTAACCTTTCATTCATAAAAAAGT
```

TATTTTTAATATCAAAGGTCACACCACCACTTTAGATATACAATTTACAGA
AATTTTAAAATATATATAAAAGCTTTTATAACAGTATTTCAAACAAAAATA
TCTAAATATAATCAGATTTAAAAAATGAATATTTATGTAAAATAATGGTTTT
ATTTTCATTATCAATCTGATATGATAGATGCATATTAATAATCATAAAAAA
5     TTATTTGAGGTGT


Sequence 3351
step.1001d08.cons.ok
10    TCCCCTTCGGGGGACAGAGTGACAGGTGGTGCATGGTTGTCGTCAGCTCGT
GTCGTGAGATGTTGGGTTAAGTCCCGCAACGAGCGCAACCCTTAAGCTTAG
TTGCCATCATTAAGTTGGGCACTCTAAGTTGACTGCCGGTGACAAACCGGA
GGAAGGTGGGGATGACGTCAAATCATCATGCCCCTTATGATTTGGGCTACA
CACGTGCTACAATGGACAATACAAAGGGCAGCGAAACCGCGAGGTCAAGCA
15    AATCCCATAAAGTTGTTCTCAGTTCGGATTGTAGTCTGCAACTCGACTATA
TGAAGCTGGAATCGCTAGTAATCGTAGATCAGCATGCTACGGTGAATACGT
TCCCGGGTCTTGTACACACCGCCCGTCACACCACGAGAGTTTGTAACACCC
GAAGCCGGTGGAGTAACCATTTGGAGCTAGCCGTCGAAGGTGGGACAAATG
ATTGGGGTGAAGTCGTAACAAGGTAGCCGTATCGGAAGGTGCGGCTGGATC
20    ACCTCCTTTCTAAGGATATATTCGGAACATCTTCTACGAAGATGAGGGAAT
AACGTGACATATTGTATTCAGTTTTGAATGTTTATTAACATTCATTTGTAC
ATTGAAAACTAGATAAGTAAGTAAGATTTTACCAAGCAAAACCGAGTGAAT
AGAGTTTTAAATAAGCTTGAATTCATAAATAATCGCTAGTGTTCGAAAGAA
CACTCACAAGATTAATAACTAGTTTTAGCTATTTATTTTGAATAACAATTC
25    AAAATATGGTGGAAACATAGATTAAGTTATTAAGGGCGCACGGTGGATGCC
TTGGCACTAGAAGCCGATGAAGGACGTTACTAACGACGATATGCTTTGGGT
AGCTGTAAGTAAGCGTTGATCCAGAGATTTCCGAATGGGGGAACCCAGCAT
GAGTTATGTCATGTTATCGATATGTGAATTTATAGCATGTCAGAAGGCAGA
CCCGGAGAACTGAAACATCTTAGTACCCGGAGGAAGAGAAAGAAAAATCGA
30    TTCCCTGAGTAGCGGCGAGCGAAACGGGAAGAGCCCAAACCAACAAGCTTG
CTTGTTGGGGTTGTAGGACACTCTATACGGAGTTACAAAAGAACATGTTAG
ACGAATCATCTGGAAAGATGAATCAAAGAAGGTAATAATCCTGTAGTCGAA
AACATATTCTCTCTTGAGTGGATCCTGAGTACGACGGAGCACGTGAAATTC
CGTCGGAATCTGGGAGGACCATCTCCTAAGGCTAAATACTCTCTAGTGACC
35    GATAGTGAACCAGTACCGTGAGGGAAAGGTGAAAAGTACCCCGGAAGGGGA
GTGAAAGAGAACTTGAAACCGTGTGCTTACAAGTAGTCAGAGCCCGTTAAT
GGGTGATGGCGTGCCTTTTGTAGAATGAACCGGCGAGTTACGATCTGATGC
AAGGTTAAGCAGCAAATGCGGAGCCGCAGCGAAAGCGAGTCTGAATAGGGC
GTTGAGTATTTGGTCGTAGACCCGAAACCAGGTGATCTACCCTTGGTCAGG
40    TTGAAGTTCAGGTAACACTGAATGGAGGACCGAACCGACTTACGTTGAAAA
GTGAGCGGATGAACTGAGGGTAGCGGAGAAATTCCAATCGAACTTGGAGAT
AGCTGGTTCTCTCCGAAATAGCTTTAGGGCTAGCCTCAAGTGATGATTATT
GGAGGTAGAGCACTGTTTGGACGAGGGGCCCCTCTCGGGTTACCGAATTCA
GACAAACTCCGAATGCCAATTAATTTAACTTGGGAGTCAGAACATGGGTGA
45    TAAGGTCCGTGTTCGAAAGGGAAACAGCCCAGACCACCAGCTAAGGTCCCA
AAATATATGTTAAGTGGAAAAGGATGTGGCGTTGCCCAGACAACTAGGATG
TTGGCTTAGAAGCAGCCATCATTTAAAGAGTGCGTAATAGCTCACTAGTCG
AGTGACACTGCGCCGAAAATGTACCGGGGCTAAACATATTACCGAAGCTGT
GGATTGTCCTTTGGACAATGGTAGGAGAGCGTTCTAAGGGCGTCGAAGCAT
50    GATCGCAAGGACATGTGGAGCGCTTAGAAGTGAGAATGCCGGTGTGAGTAG
CGAAAGACGGGTGAGAATCCCGTCCACCGATTGACTAAGGTTTCCAGAGGA
AGGCTCGTCCGCTCTGGGTTAGTCGGGTCCTAAGCTGAGGCCGACAGGCGT
AGGCGATGGATAACAGGTTGATATTCCTGTACCACCTAGTATCGTTTTAAT
CGATGGGGGGACGCAGTAGGATAGGCGAAGCGTGCTGTTGGAGTGCACGTC
55    CAAGCAGTAAGGCTGAGTGTTAGGCAAATCCGGCACTCATAAGGCTGAGCT
GTGATGGGGAGAGGAAATTGTTTCCTCGAGTCGTTGATTTCACACTGCCGA
GAAAAGCCTCTAGATAGATAACAGGTGCCCGTACCGCAAACCGACACAGGT
AGTCAAGATGAGAATTCTAAGGTGAGCGAGCGAACTCTCGTTAAGGAACTC
GGCAAAATGACCCCGTAACTTCGGGAGAAGGGGTGCTCTTTAGGGTTCACG

CCCAGAAGAGCCGCAGTGAATAGGCCCAAGCGACTGTTTATCAAAAACACA
GGTCTCTGCTAAACCGTAAGGTGATGTATAGGGGCTGACGCCTGCCCGGTG
CTGGAAGGTTAAGAGGAGTGGTTAGCTTCTGCGAAGCTACGAATCGAAGCC
CCAGTAAACGGCGGCCGTAACTATAACGGTCCTAAGGTAGCGAAATTCCTT
5    GTCGGGTAAGTTCCGACCCGCACGAAAGGCGTAACGATTTGGGCACTGTCT
CAACGAGAGACTCGGTGAAATCATAGTACCTGTGAAGATGCAGGTTACCCG
CGACAGGACGGAAAGACCCCGTGGAGCTTTACTGTAGCCTGATATTGAAAT
TCGGCACAGCTTGTACAGGATAGGTAGGAGCCTTTGAAACGTGAGCGCTAG
CTTACGTGGAGGCGTTGGTGGGATACTACCCTAGCTGTGTTGGCTTTCTAA
10   CCCGCACCACTTATCGTGGTGGGAGACAGTGTCAGGCGGGCAGTTTGACTG
GGGCGGTCGCCTCCTAAAAGGTAACGGAGGCGCTCAAAGGTTCCCTCAGAA
TGGTTGGAAATCATTCATAGAGTGTAAAGGCATAAGGGAGCTTGACTGCGA
GACCTACAAGTCGAGCAGGGTCGAAAGACGGACTTAGTGATCCGGTGGTTC
CGCATGGAAGGGCCATCGCTCAACGGATAAAAGCTACCCCGGGGATAACAG
15   GCTTATCTCCCCCAAGAGTTCACATCGACGGGGAGGTTTGGCACCTCGATG
TCGGCTCATCGCATCCTGGGGCTGTAGTCGGTCCCAAGGGTTGGGCTGTTC
GCCCATTAAAGCGGTACGCGAGCTGGGTTCAGAACGTCGTGAGACAGTTCG
GTCCCTATCCGTCGTGGGCGTAGGAAATTTGAGAGGAGCTGTCCTTAGTAC
GAGAGGACCGGGATGGACATACCTCTGGTGTACCAGTTGTCGTGC
20


Sequence 3352
step.1001e05.cons.ok
TTCATTACCTAATGCAGGTAAAATTTTCTTAATTGCCTGTGGCATAACTAC
25   TGATTGCATAGTCTGTCTATAATTCAATCCTAAACTTCTAGCAGCTTCAGT
TTGTCCTTTATCTACAGCATTAATACCAGCTCTAATAATCTCGGCAATGTA
GGCTGAAGAATTTATAACAAGCGCAATAGTACCACAAATTAAAGCTGAAAT
ATCTAACCCCAAAGCTGCAGTAGTCCCAAAAAATACAATAAACACTTGCAC
AAGCATAGGAGTACCTCTTAAAAATTCTATGTATATACTTGCTATCCATTG
30   TAATGGTCTAATTTTACTTATTTTTAATAAAGCTATAAAAGAGCCTAAAAT
AGACCCCAATACCACACCTACTAATGAAATCAAAATAGTATTTTTAATTCC
TTTAATGAAAAAACTGCCATATTTAGAAATAAAATTTCCATCATCTTGCAT
ATCTTCGGCTGCTTTTGTCATGTATTGATCAATTAAATTTTTTTCTTTTAC
ATTATCAATCGTTTGATTTAGCTTATCTAAAAGTACAGGAGAGTTCTTAGG
35   AACAGCTATACATGTTTGTTTCTTTTCTTCATTAAATTTAATCTTTGAAAA
TGTTAGTTCTGAATTTTGTTTTAAATATGCCTCACCTACTGGTTTTTCAAC
AACTACACCTGCTACTTTTCCACTTTTTAAAGATAATATAGCTTCAGGTAA
TCTATTGAGTGAAGAAATTTTACTATCTTCAATCTCAGTTTGTGCAATTTT
TTCTTGATCTGTCCCTTTTTGCGCAGCAATTTTTTTGCCTTCAAAATCTTT
40   ACCTGTTGTCATCATCATTCCAGGATGTGTGAACATTACTTTTCGTGCTTC
ATTAAGCATCATCCCCCATTCGGCTGTAGGTGCCTTAACACCTAATCCAAG
GAACGAGAATCCTGACATTTGTAATATCATTGAGCACATCGAACTACTAGC
AATAATCGCTATGTCAGTAAAGGTTAGTGGCAAAATATGTTTGCGAATGAT
TGTTAAATCATTCATACCAATTACTTTGGCAAATTTTACATGATCAGCTTC
45   AATATATTGCATTACACTGGTTCGAATCACGCGACAAAACCACGCCCATCG
AGTCAATATAAACGCAATAATAATATTTTCTACACCCATGCCAAACAACGT
AATCAATGCCAATGTCACCCACATAGCTTGGAAAAGCTAACATCACATCGCA
TATACGCATAATTATTGCATCGATATAACCTGGGAAATAACCTGAAATAAA
CCCAAGTATCGCTCCTATCACAACGGAAATAATCAATGCGACAAATACATA
50.  TAACAAACTAGGTCTTATGGCGTATATTATCCGTGTTAATACGTCCCGGCC
TAAATGGTCTGTCCCCAACCAGTGAGACCAACTTATACCAGCAAATTTATT
TGCTGTATCAATGTGATTCGCTTCATAAAATGTAATTAAAGGAGCTAAAAC
TCCAAGCACTACGTAGATTGTAATAATAGCTATAGCAACTATGGCACCTCT
ATCTTGAAATAGACGACGTAACACAATCATTTCATTGCCTCCCTTAAACGC
55   GGATTTAATATTGCATTAATGATATCTGCAAGTGTATTAAAAAGAATAAAC
AACACCGCGACAATGAGCACGTATGCTTGAATTACTGGAAAATCATGTTCA
GTAATCGCTTTAACACTTAATGGTCCTAAACCTGGCCACGCAAATACATTT
TCAATAACGACTAAACCTCCCATGATCATTGGTATCGACATACAAAAAATA
GATACGGCTACCTGTAATGCATTACGTAATACATGCAGCAACATTGTCAGA


855

CGAGGTACACCACTAGCTTTTAAATACAAAACATAATCTTCATTTAATTGT
TCAAGCATAGATCGTCGGACATTTCTAAAATAAATTCCTGCATAGGTAATC
GTAATAACAAAGACTGGAAGAATATAACTTTCTGGACCTGTGAGTCCTGAA
GTTGGTAATAAATTGAGTTTTACAGATACATAAATAACTAATATTGAAGCT

5    ACCCAATATGAAGGCAGAGCTGTTAGTAGAAAAGCTATTGAACGTATAGCA
CGATCTGTAAATCGACCACTAGTCAAAGCACTCACTATGCCTAGAACAACA
GATGTTATCATCACCACCACACTGGAAATCAAAGTGAGTTTTAATGTATTC
ATAAACGCTGGACCCATACGCTCAGATACTGGATCCCCTGTGATATAACTC
GTTCCAAATTTAAACTGCAAGGCATCAACAAGCCAATTTTTATATTGCGTG

10   AATATCGGTTGATTTAATCCATATTTTTCTTTAGTTTCAGCTATTAAATCT
TCTGTTATATTAGGTGTTCCTTGGGCATGTAAAATTGTTACTGCTGGATCT
TCGTTTGTGAGATAGGTCATCATGAAAGTCATAAAAGTCACTACAATTAAT
AATGGAATCATGAGCGTAAGACGTTTCATAATAAATTTAACCATGACGACG
CTCCTTTATTTATATTGCATTTCATTAAATGGAAGCTCATATTGTGATTGT

15   GTAAAGGATAATTTTTGTAAATCTTTAGGTGCAACAACAGTCATTCTGCCG
TGAGTAATTGGAATAAAGACACCTTCATCGTCAACTTGTTTTAAAATTTGC
TGATACGCTTTTGAACGTGCCTTTGCATCTTGAATTTTGAAAGCATCATCA
ATATCTTTGTACAGTTGTGTCTTATTCTTAATACCTAATGTTGCACTTTCA
TATCCCGTTTTCGTCTTAAATGCTGCTAATGTACTCTGAGGATCATATAAT

20   AGCCCCCATGTTTGATTAAACATTAAATCGTAATCACCAGAACTACGACGT
TCAGCTACTTTATCTGACGTTTCTCCGTTTATATCTAGTTGTACACCTAAC
TTCTTAAATTCTGCTTCTAAGAATTCAGCTTGTTCTTTTTGACTGGAAGAC
CCTTTGTCATAATACAACTTAAGATTCAAATCTTTGCCCTCTTTTTGACGA
ACTTGTCGATCTTTTGAAAGCACCCATCCAGCCTTGTCTAATAACGCTTGC

25   GCTTTTTTCTTATCATATGTTCTTGTTGGTAAATTAAAGTTTATATCTGTC
ACATTTTTAGCAAATAGTTGTGTGGCTGGCTTTTCTTGTTTGTCTAAAATA
TCTTGAGCTATTTTATCTCTATTTACCATGTGACCTAATGCTTGTCTGACT
GCTTTATCACTGACTGCACTATCTTTCTTACCAGAATTAACAACAAGCATT
TTTGTATTCATAGCCTGGCTACGTTTTACTTGGTAGCTTCCGGTTTCTTTT

30   AATTGTTTTAAACTATCTTTATCTAAGCTGTCTGTACCTCTATCATCTGTA
TAAGCAAAGTTGGTTTCTCCTTTTTTCATTGATAAAAATGTTGTTTCTCCT
GCAGGTTTAACTTTTGCTTCTACTTTATTTAATTTTGCCTTTTCTCCCCAA
TATTGATTATTTTTATTAAACTCTGCAGATATATCTTTTTTGTGTC

35

Sequence 3353
step.1001e06.cons.ok
TTAAAGTAGGTCAACCAAGAGAAGAAGGAACTCAAGTAGGTCCTATTATTA
GTAAAAAACAATTTGATCAAGTTCAGGATTACATTGATAAAGGTATCAATG

40   AAGGAGCAGAGTTATTCTACGGTGGTCCAGGTAAACCAGAAGGATTAGATA
AAGGATATTTTGCGCGTCCAACTATATTTATTAATGTAGATAATCACATGA
CGATTGCACAAGAAGAAATTTTCGGACCTGTCATGTCTGTAATCACTTATA
ATAATCTTGACGAAGCTATTGAAATTGCTAATGATACTAAATATGGTCTTG
CTGGCTATGTTATTGGTAAAGATAAAGATACATTGAGACATGTTGCTCGTT

45   CAATCGAAGCTGGTACAATTGAAATTAATGAAGCTGGTAGAAAACCAGACT
TACCATTCGGTGGTTATAAAGAATCTGGCTTAGGTCGTGAATGGGGCGATT
ATGGTATCGAAGAATTCTTAGAAGTAAAATCAATAGCTGGATATTTTAAAT
AATTAGTAAATTAAATTAAATGTTGCAATTAAACAAAGCACGGAACTAATT
TTGTTCTGTGTTTTGTTTTTTATGTTAAAATTCCTAAAATCATTTTGTAAT

50   TAAAATGAAAATTTGCTATAGTTAGTCTAGTAATTAGAAGGTTGAGAAATA
TACAAACGTATAAATATAGCTATAATGTGACTTTTTATTATTGTCAGTCAC
ATAAATATCGAGATAAGAATCTATTATTTGGTTAAGCTAACTATACGTTTA
TTATTTTTGAACTTCGTAAAATAAATAATTAATGAAATGGGTGCAAGTCTA
TGCCTGGTAGAGAACGAACATCTCCGCAGTATGAATCTTTCCACGAATTAT

55   ATAAAAACAATACTACTAAAGAGCTCACTCAAAAAGCCAAATCTTTAAAAT
TAACTAATTACAGTAAATTAAATAAAAAAGAATTGGTACTTGCCATTATGG
AAGCACAAATGGAAAAGATGGTAATTACTATATGGAAGGAATATTAGATG
ATATTCAACAAGATGGTATATGGTTTCTTAAGAACCGTTAACTATTCTAAAG
GTGAGAAGGATATTTATATTTCTGCAAGCCAAATTCGACGTTTTGAAATAA

```
      AACGTGGTGATAAAGTAACGGGTAAAGTTCGTAAACCAAAAGATAATGAAA
      AATATTATGGTCTACTTCAAGTTGATTTTGTAAACGACCATAATGCAGAAG
      AAGTCAAAAAACGTCCTTCACTTCCAAGCTTTAACACCTCTTTATCCGGAA
      GAAAGAATCCTATTAGAAACGCAATCTACAAATTATTCCACTCGTATTATG
 5    GATTTAGTCACACCAATAGGTCTTGGTCAACGTGGTCTTATAGTTGCACCA
      CCTAAAGCTGGTAAGACAAGTTTATTAAAAGAAATCGCTAACGCAATAGCG
      AGTAATAAACCGGAAGCGAATTTGTTTATATTACTAGTAGACCTACTCAGG
      AATTAAGAGAGCAAGGTGTACGTGTGAAACTGTCGGCTGTTAAGGATATTG
      TTGATGGTAAAGATATCGTACTTGTAGATGATTCGATTGTTCGAGGTACAA
10    CGATTAAACGCATAGTTAAAATGCTTAAGGATTCAGGAGCTAACCGCATTC
      ACGTAAGAATTGCTTCTCCCGAATTCATGTTCCCTAGTTTTTATGGTATTG
      ACGTATCTACAACAGCTGAACTCATCTCAGCAAGTAAGTCTCCTGAGGAAA
      TTAAAAATCATATTGGTGCAGATTCTCTTGCTTATTTAAGCGTTGATGGCT
      TAATCGAGTCTATAGGACTTGATTATGATGCGCCATATCATGGCTTGTGTG
15    TAGAAAGTTTTACAGGTGATTATCCAGCAGGACTTTACGATTATGAGAAAA
      ATTATAAAAAGCATTTAAGTGAACGTCAAAAATCATATATAGCTAATAATA
      AACATTATTTTGATAGTGAGGGAAATTTACATGTCTAAAGCATATGAGGAA
      TCTGGTGTAAATATTCAAGCAGGATATGAAGCAGTCGAAAGAATAACAAGT
      CATGTTGAACGTACATTGCGCAAAGAAGTATTAGGTGGTTTAGGTGGATTT
20    GGTGCAACATTTGATTTGTCTCAATTAAAAATGAAAGCGCCAGTTCTGGTA
      TCAGGTACTGACGGTGTGGGTACAAAGTTAAAATTAGCAATTGACTATGGA
      AAGCATGACACAATTGGTATTGATGCTGTCGCAATGTGTGTAAATGATATT
      TTAACAACAGGTGCTGAACCTTTATACTTTTTAGACTATATTGCCACGAAT
      AAAGTAGTGCCAAGTACTATAGAGCAAATCGTTAAAGGTATAAGTGACGGT
25    TGCGAACAAACCAATACGGCACTTATAGGCGGTGAAACTGCTGAAATGGGA
      GAAATGTATCATGAAGGTGAATATGATATTGCTGGTTTTGCAGTAGGAGCG
      GTAGAGAAAGAGGACTATATTGATGGTTCAAATGTTGAAGAAGGACAAGCA
      ATTATTGGTTTAGCTTCAAGTGGTATTCATTCAAATGGCTATAGTCTAGTT
      AGAAAAATGATAAAAGAATCAGGAGTTCAATTACATGATCAATTTAATGGT
30    CAAACCTTTTTAGAAACCTTCCTTGCACCAACAAAATTGTATGTAAAGCCT
      ATTCTTGAATTAAAGAAACATATTGATATCAAAGCGATGAGCCATATTACT
      GGTGGAGGTTTCTATGAAAATATTCCGCGTGCCCTTCCTAAAGGTTTATCA
      GCAAAAATAGATACACAATCATTCCCAACGTTGGAAGTCTTTAATTGGCTT
      CAAAAACAGGGCAACATTTCAACGAATGAAATGTATAACATATTTAATATG
35    GGTATTGGATATACAATTATTGTTGACAAAAAAGATGTTCAAACAACATTA
      ACAACGTTACGTGCAATGGATACAACTGCATATGAAATTGGTGAGATTATA
      AAAGATGATGATACACCTATTCATTTATTGGAGGTAGAATAGTGACCAATA
      TAGCGATTTTTTGCCTCAGGATCAGGTAGTAACTTTGAAAATATTGTAAAAC
      ATATCCAAACAGGGCAATTATCTGGTATCAATGTAACGGCACTGTATACAG
40    ATAATGAAGGGGTACCCTGTATTGATAGAGCTAAAAATTTAAATATTCCAA
      TTCATATTAACAAGCCAAAAGACTTTTCATCCAAATCTTTATATGAACAGC
      ATCTACTTAAATTATTATCCAGTGAGGAAGTTCAGTGGATTGTATTAGCTG
      GTTATATGAGACTAGTTGGCCAAGACTTATTACAAGCTTATGAAGGACGGA
      TATTAAATATACATCCCTCATTATTGCCTAAATTCAAAGGTTTAGATGCCA
45    TAGGTCAAGCGTTAGAAAGTGGAGATACTGTAACTGGATCAACTGTCCATT
      ATGTAGATAGTGGGATGGATACGGGAGAAATTATTGAACAACAGCAATGTG
      ATATAAAACCGGACGATACTAAAGAACAATTAGAAGATAGAGTAAAACATT
      TAGAATATGAACTTTATCCAAGAGTTATAGCTAAAATCATTAAATAGAGGA
      GTATCTTATAGAATGAAAA
50
```

```
      Sequence 3354
      step.1001e12.cons.ok
      TAGTGAATGAAGACCCTGAACGTAAAATCATTATCGTTTCAGCTCCAGGCA
55    AAAGGCATAATGACGACATTAAAACTACTGATTTATTAATTCGTCTCTATG
      AAAAAGTACTTAATAAATTAAATTATGAAAGTAAAAAACAAGAAATTATCC
      AAAGATATGCTGATATAGTAGAAGAATTAGGTATAGGAAATGACATTTTAA
      TAACAATTAATGACACTTTAGAGGAATACATTAAACATCTTTCTGACAAAC
      CTAACCGTTTATATGATGCTTTATTATCTTGTGGCGAAAATTTTAATGCTC
```

```
      AATTAATAGCCCAGTATAATAATAGTCAAGGTATTCCTACTCGTTATATTT
      CTCCTAAAGAAGCTGGATTAACTGTAACTGATTTACCACAGCAAGCTCAAA
      TTTTAGATTCCGCATATAATGAAATATACAAATTGCGTGATTATGATGAAA
      AGCTAATTATTCCTGGTTTTTTCGGAGTTTCAAAGCAAAATTATATCGTTA
   5  CGTTTCCACGCGGTGGTTCTGACATAACTGGTGCTATCATAGCACGTGGCG
      TCCGAGCCTCACTTTATGAGAACTTCACTGATGTATCAGGAATATATAAAG
      CTAATCCGAATATCATAAATAATCCTGAACTCATAGAGGAAATAACTTATA
      GAGAAATGCGAGAGCTATCTTATGCAGGATTTGGAGTTTTTCACGATGAAG
      CTCTACAACCTTTATACAAAGATCGAATTCCCGTAGTTATCAAAAATACTA
  10  ATCGTCCAAATGATAAAGGGACCTACATTTTACATGACCGTGAAATCGATT
      CTAAAAATGTCATTAGTGGAATTAGTTGTGATAAAGGCTTTACTGTGATTA
      ATATTAAAAAATATTTAATGAATAGATTAGTTGGATTTACACGAAAGATTC
      TTGGCGTTTTAGAAGAATTTAATATATCATTTGACCACATGCCTTCTGGTA
      TTGATAACATAAGTATTATCATGCGTACAAATCAAATTCAAGGTAAAGAAA
  15  GTCAAGTTCTTAATGCCATACGCAAACGTTGTGAAGTTGATGAATTAAGTA
      TCGACCATGATTTAGCAGTACTAATGATTGTTGGTGAAGGTATGAATCAAG
      TTGTTGGTACAGCTAGTAAAATTACTCACGCCCTTTCAGAATCAAACATTA
      ATTTAATAATGATTAACCAAGGTGCTTCTGAAATTTCAATGATGTTTGGAA
      TTCATGAAGCAGATGCTGAAAAAGCAGTATTATCTACGTACGAATTTTGTT
  20  ACAACGGTGTTTGTTTAAAAAAATTTGTGTAAATAAAATAAAAAAGATTTCT
      ACATCACTTTCAATACATAAAGTAAATTTATTAACTCTGTATCAATGATGT
      AGAAATCTTTTTAAGATATCTTTTTAGTGGAACTATTTACTATAGCTTTTT
      TAATACGCTGTTCACCTACTACCAATTTTAAAATATAGAATTGATAATAAC
      TAAGAGAATCAATAACAGCTTTATAATTACTAAACTGATGATATTTTTGTT
  25  CACTCATTAAATCATACATAAGAACTATTTCTGGCTTAATATAATCTACGT
      TCCAACTTAAAGAATGAAAATATATATTATTTTTTGGTAAACGTATATGAT
      GATCTAACCTAAATAGCCATTCATCATTTACAACATCATAAACAAATATTG
      TCATGACTTCTGTATCGTTTTTAAAGACTTTAACATGTTCAACATCTGATA
      AAGATAAAGATTCTCTTAATCTTTCATATAACTGATTATCATAATAGTGAA
  30  TCGTAAATTCTTCAGGAATCACCCTGATAATTTCCAATAACTTTTTGCGCT
      CAACACAAATATCTATTTTCGATGGTAATATAAAATCATCATTCAGAAATA
      GATGCATCGCTACCTCGCCGTGAAATTTGAATAAATGAGATGATGTTTTCC
      TTAAAATTTCGACTATTTTATTCACTGGTTTTTTCTTCTTTAACAAACATGT
      TACCACCCCTATAGCAAGCGCTTTCACAATCAAATTATACATTTTCTTTTT
  35  TCACTTATCAAACAAATAAATGTGAACATTTAGTAAAATTTATTTTATTTT
      CAATAAAATTATTTTCATAAAAAGTTACTATATCACTATTCGCTACGCATA
      TGATAGGGCTCTCATAAAAAATTCACGAAAGTTTAATTGATATGTAAATATT
      TCGTCACTACTATTTGATAAAACACATTTCCCCTTTCATTTTCATTGAACT
      CAAAATTTAAGAAGTTTATAATTTATCTTGCAAATAGAATTTAATATTAAA
  40  AGATTTGGGTGAGTAGTAATGTCTTTAACGATTATCTTATCTATAATAATT
      ATTATTTTGATTATGGCTATGGTTCTTAATCAAAAATTTATGAAAGATAGG
      GTTGAAACAGAAGAATATGCTAGAAATCAATTAATCTCTAAAAATTCAATT
      TTAAGTGAAGAAAATTTATCATTGAAAAACCAAATGTTAAGTACAAACAAT
      GACGTCGGTCAACACGCTTTTAAAAACGCCAAGCGTGAATTAAGAAAAATA
  45  TTAAATAGATTTAAAGAAGAGGGTCGTTTACGATCATATACAATTGTTCCT
      ACGAGTAATTTGGCTGTTAAACATCCCCTTTTCGAATATGCACGTTCATTC
      GATTTTATTATCATTACTGATGTTGGTTTGATAAATGTGGATGTTAAAAAT
      TGGAACCAAAAAACGTTTTATCATTTTGATGTGCCAGATCAACATCTTGAA
      GAAGGACAACCACAATATAATACCGAAAAGTTGTCGGTCATTATATTAGC
  50  AATCGATATCATAGTCAGTTTAAAACAACACGTTCTGGTGTCTATACTTTT
      ATTGAGATTTTACAGGATAATCGTGTAATATATGAATTTTATGACCACGAT
      CCATACGATAAAGCCGCAAACAATGCAAAGCATTAAAAGATAAAATTGAA
      AATGATTATAATTTTAAAATTCAAAGTATTGGCGTCATATATTTTAGTGAT
      GGTAGCGTTAATATTATTGAAGGATCCGACGAGAGTGATAAATACGTCGAC
  55  ACCGTATCTACACCGATATCACTTGAAAAAGTAATTGAGGAAGCTATCGAT
      TTATCTAAGCACCCCCTTACTGATAAACAAATCGAAGAAATTTCTGAAAAC
      TTTAAACAACATATGAATAATTAAAATTCAAAAGAGTTTGGGGCATGAAG
      TTATTTGACTAATAAAAAGACAATTTCTATTTATAAAATGTAGAAATTGTC
      TTTTTATTATAATTATTTAAATTGTTGAGCAACTACTTTTTTTATATTCAT
```

```
TAGCATTAATACAAATGCAATTTCTCTTTTGACTTTAATTATGTTTCGAAT
CGACATTCTGATGAAACTTAAAATAACTTTCGTAAATCCATCAACAAGTTC
CACATCAAATTTCTTTAACTGCAAACGTTTCTTCTAAAATAAGAAATCTTA
TCCCAATTTGAATTTTAAAATATCTTAAATTGTAATTTTACATTCCAACTT
ATTAAATTATCTATTATCAACAATGTTGTCATATATATTTTTAATTTTAAC
GTTTTAATACGATTCAGTAGTCATAGCACACATTTATATGACAACACTTAT
AGCAATTATTTCATGAATAAATAAACTTAATTATAAACTGTATCAGTTTAC
ATTAGAAATCTTTTTAAAACATAATGTTTTTCACTAGACTAAGAAGTAGAT
ATCCTAATTTCTACTTTATTGAAACACTAGTCAGAAACGTAGTTGAAATCT
ACTAAATATCACTCTTTTTTTATTCTAATCTCTTTAGCTAGAATAAGAAAA
CGCTGCTAAACAAAATGGTCCAACTGAGTTTCTACTTATTGATATAACAAA
TGAATTTTATATTTCCACTCAATATCAGACTCTAAATTATCGCTACTAGCT
CCAAATATTTAATTTTCTCTTTTTTGCCTCTTCTTGAGCGTTACGAAGGT
ATCCCTATATTTTCCGTTTGGAGAAAAATACTTTTCTCTGGCTAGACCTTT
TTTCACCAGCGATT
```

Sequence 3355
step.1001f07.cons.ok

```
ACTTGAAAAATCAAATGCAGATCGTGACATTTTCTATTTAAAATCTATTGA
TAACAATATTCGTGAATATCATATAGCAGAATAATCTAAACTCAATTCTTT
TGCTTTTTTAATTTTTTAAAATGAAAGAAACGAATTTACTATTCATACTTG
CCCCCTAAAAACAACGTCATATAGCTGTTTTTAGGGGGTTTTTAATCCCAA
GATAGTGCATTTCATCTATTATAGACCTCTACTATTTATATACTGTTTACC
TACTATTCATAGATAAAAACGAATCCTATCCATAGATCAACATTGACAAAT
AACTAAAGCAATTCTATACAATAACGTATCTCAAGGTTTCGTTGTTTTTTT
AGGTCTGTATAAATGTGTCATATCTATTTCTTCCAATTTTAATAATTTTAT
TTTGTTGTCTAAAGTACCTCCGTATCCTGTTAAATTTCCATCTTTACCAAC
CACTCGATGACATGGTATAAGTATAGAAATAGGATTACTCCCAACAGCACC
TCCTACAGCCTGAGCAGACATTGCTGGCTTATTCATTGCTCGTCCTACTGT
TTTAGCAATCTCTCCATATGTTTTAAGTTCTCTATAACGAATTGTCTGAAG
TTCATTCCATACCTTCTGTTGAAAATCAGTGCCTTTTGGTGCTAACGGTAC
TTTTATCTCAGGATAATTCCCTTTAAAATACTCTCTTAACCACGCTTTAGT
ATCTTTAAAAACATCTAACGTATCCTTTGTTTTAATTCCTTTCAAATGTTG
ATTAGGGTATAAGACATGTGACAACGAGATGCCATCGTTTGTAATTAATTG
CAATAAGCCAACGGGCGAGTCATACATTGATTGATACATATTTCTCCAACT
CCTTAAATGTGTATATCATTCTAGTCATATCCTTATAAACTAACGCATTCT
ATTCTGAACATAATTTTATAATATAGATAATTGCTATATTTATGAAATTGG
ATAAACATTCGAATAGAATGTGATACACCATTAACGATAACTAAAATTGAA
AAATGTCATCATTTATATATTAAGTCAATTATCGACAGGTCATACTTTAAG
AAAGTCTGGAACTTTGATTTACGTCCCAGACTGGTCAAGATATAGAATTTA
AAGTCCAACTCACCCTTTTGATTAGATGTCATTATATTTATTGAAAATTAT
AAAGTTCATACCCTTTATTTTTTGCATATTCTAATGCGAGAATTTCTAAAA
CATTAATGAACGCAGGTTCGTCTTTTACTTTTATATATTGAAAAGTGAGGC
CGTATTTATCTGCAGTAATGTCTTTTAAATGATTTGGATTTAAGTCCTTAT
TTTGAGCATAAGCTTTTCTAAATTTAGTCGTACCTGTTTCTGAACGTGAAT
TGTATTGTCCTAATTTTGTATACAAATCACTATTTATAGGTTGATGTGTGG
ATGAATATCCTCCAGTTTGACCTACAAACACTGGTTTTTTCTCATCGAAAA
TCACATAGACACCAAACTTATCCTTAACTTTTCCCCTATCCTCTGGTGTAA
ACGCCAAAGCCTTTTTATGATCGTACTTAAATAATTTTTTTAGTTTCTTCT
GATTATCTTCTAGTAACTCATTTAATTTTTTGCTACCCTTCATACATTTAT
TCCTCCCCATTAGCAATAAACTTATATACATACATTTTACATGTAGTATAT
CACTTTTTAAAACTTTTAATCATTTTCAGTTTTAATAATTTTCACTATTAC
CTCAGGTACTAAAAAAACCTAAGATTGAAAACACAAAACCATTTAAAACGTA
GCCTAAAATACTCCATGTGTGATTATAGCTCATCTGTAATTGTGTTCTCGT
TTGAGCTATACGGTCACGTTCAAATCCATGCACGAGTCCCGCAACTACTGC
TGCAATAATCCCTGATGCATGGAATAATTCAGCAATCAAATAAGTGACAAA
CGGCGTCAGCAATTGTATAAATGTAAACATGTTAATGTTCTCAATGCCACG
TCTCATTAATGTTAAACGAAATCTCACTAAAGCCATGCCTATAATTAGACC
```

```
       AACAATAGCACCTCCAATAGAGGCAATCAAAAATTGTTGAATCGCATCAAA
       AATGGAAAAAGTTCCAGTAATCAATACACCTACGGCAATTTTAAATGAAAT
       GATACCTGCAGCATCATTAAGTAAAGATTCACCTTCAAGAATGGTCATTGA
       ACCTTTTGGTAAAACCTTACCTTTCGTGATTGCTTGTACCGCAACAGCGTC
  5    TGTAGGACAAAGTATAGCCGCAATTGCGAAAGCAGCACCTATAGGTAATTC
       AGGCCAAATCCAATGAATAAACAATCCTACGCCTATAACGGTAGTTATGAC
       AAGACCTAGAGCCATCATCATAACAGGTTTGATATACTTCCTTAAATGTAC
       TCTTGAAACATTAACACCCTCAACGAATAGTAGTGGTGCAATCAATGTAAC
       CATAAATAATTCAGAATCAAAATTGAATTCAACTGGAATGGGAGTGAGGTA
  10   TAAAATCATCCCAAGGAAAATTTGGATAAATGCAAGAGGTACTTTAGGTAT
       GAAAGTATGTACAAAGGAGCTCACAATAACTAGTGCGACAAATATGAGTAG
       AGTTTCAAATATCTCCACACTTTCACCTCTTTATTAATTTTTAATTGCAAA
       TTTAGTCATCAAAACAATAAAAAGTTGCAAGGTTATCTACTGCAACACTGA
       AGAAGTGAAATAATAACAATAGAGTCATTCCTGTAAGGACATAACTATTAA
  15   CACTCAATCATGCCACAAATATACTATTCAAAACTCAATTATTTTCTTTAT
       CAGTGAAATCAAACATGAACATTTAAATTAACTTATGAAGCTCGCTTCTTC
       GAACGAGAAAATGCTTATGTGTAATTTCATTGAAAAAATTAGTGACAACTG
       TCATTCACAATTATTCTTGTTTACAATTTTATCATTTATTAACCAGATTCT
       TAAATAAATAGCTTTTATATCATTAATATCTTCACCTCAGATATACAAAAA
  20   ACTACTAGCATATAAAAGATATATCTTTTATATGCTAGTAGTTTAAAAGCT
       TAAACAATCAATCGTCCATTTTATTAATATATTCTTTTTCTTTATAGTATA
       ATTTTTCTCGTCTTTCTTCTCGTAGTTTTGCACGTTTTTTTAATTCTTCTG
       GCGTATTAAACTCTTTTTGGCTTAATAAGATAGTAGATTGTCTAGATGCAT
       CGTCTTTCCTGTACATATATGCACCATTACGGTAGGCCGCTCTACTTATCA
  25   AATGCATTCCAACCGGAGATGTAAGATTGATAAAAACTAGTGATAATAATA
       ATCTGACACTGAAAAAACCTGAAACTATGCACTCATACATCATAACAAATT
       TGAATCTTTTATAGATGATTTCAACCTTACTGTAAATCAAGAGATGAAATG
       GGCAAT
```

```
  30


       Sequence 3356
       step.1001g01.cons.ok
       TGGGCTGTTTCCCTTTCGAACACGGACCTTATCACCCATGTTCTGACTCCC
       AAGTTAAATTAATTGGCATTCGGAGTTTGTCTGAATTCGGTAACCCGAGAG
  35   GGGCCCCTCGTCCAAACAGTGCTCTACCTCCAATAATCATCACTTGAGGCT
       AGCCCTAAAGCTATTTCGGAGAGAACCAGCTATCTCCAAGTTCGATTGGAA
       TTTCTCCGCTACCCTCAGTTCATCCGCTCACTTTTCAACGTAAGTCGGTTC
       GGTCCTCCATTCAGTGTTACCTGAACTTCAACCTGACCAAGGGTAGATCAC
       CTGGTTTCGGGTCTACGACCAAATACTCAACGCCCTATTCAGACTCGCTTT
  40   CGCTGCGGCTCCACATTTGCTGCTTAACCTTGCATCAGATCGTAACTCGCC
       GGTTCATTCTACAAAAGGCACGCCATCACCCATTAACGGGCTCTGACTACT
       TGTAAGCACACGGTTTCAAGTTCTCTTTCACTCCCCTTCCGGGGTACTTTT
       CACCTTTCCCTCACGGTACTGGTTCACTATCGGTCACTAGAGAGTATTTAG
       CCTTAGGAGATGGTCCTCCCAGATTCCGACGGAATTTCACGTGCTCCGTCG
  45   TACTCAGGATCCACTCAAGAGAGAATATGTTTTCGACTACAGGATTATTAC
       CTTCTTTGATTCATCTTTCCAGATGATTCGTCTAACATGTTCTTTTGTAAC
       TCCGTATAGAGTGTCCTACAACCCCAACAAGCAAGCTTGTTGGTTTGGGCT
       ATTCCCGTTTCGCTCGCCGCTACTCAGGGAATCGATTTTTCTTTCTCTTCC
       TCCGGGTACTAAGATGTTTCAGTTCTCCGGGTCTGCCTTCTGACATGCTAT
  50   GAATTCACATGTCGATAACATGACATAACTCATGCTGGGTTCCCCCATTCG
       GAAATCTCTGGATCAACGCTTACTTACAGCTACCCAAAGCATATCGTCGTT
       AGTAACGTCCTTCATCGGCTTCTAGTGCCAAGGCATCCACCGTGCGCCCTT
       AATAACTTAATCTATGTTTCCACCATATTTTGAATTGTTATTCAAAATAAA
       TAGCTAAAACTAGTTATTAATCTTGTGAGTGTTCTTTCGAACACTAGCGAT
  55   TATTTATGAATTCAAGCTTATTTAAAACTCTATTCACTCGGTTTTGCTTGG
       TAAAATCTTACTTACTTATCTAGTTTTCAATGTACAAAGAATGTTAATAAA
       CATTCAAAACTGAATACAATATGTCACGTTATCCCTCATCTTCGTAGAAGA
       TGTTCCGAATATATCCTTAGAAAGGAGGTGATCCAGCCGCACCTTCCGATA
       CGGCTACCTTGTTACGACTTCACCCCAATCATTTGTCCCACCTTCGACGGC
```

TAGCTCCAAATGGTTACTCCACCGGCTTCGGGTGTTACAAACTCTCGTGGT
GTGACGGGCGGTGTGTACAAGACCCGGGAACGTATTCACCGTAGCATGCTG
ATCTACGATTACTAGCGATTCCAGCTTCATATAGTCGAGTTGCAGACTACA
ATCCGAACTGAGAACAACTTTATGGGATTTGCTTGACCTCGCGGTTTCGCT

5 GCCCTTTGTATTGTCCATTGTAGCACGTGTGTAGCCCAAATCATAAGGGGC
ATGATGATTTGACGTCATCCCCACCTTCCTCCGGTTTGTCACCGGCAGTCA
ACTTAGAGTGCCCAACTTAATGATGGCAACTAAGCTTAAGGGTTGCGCTCG
TTGCGGGACTTAACCCAACATCTCACGACACGAGCTGACGACAACCATGCA
CCACCTGTTACTCTGTCCCCCGAAGGGGAAAACTCTATCTCTAGAGGGGTC

10 AGAGGATGTCAAGATTTGGTAAGGTTCTTCGCGTTGCTTCGAATTAAACCA
CATGCTCCACCGCTTGTGCGGGTCCCCGTCAATTCCTTTGAGTTTCAACCT
TGCGGTCGTACTCCCCAGGCGGAGTGCTTAATGCGTTAGCTGCAGCACTAA
GGGGCGGAAACCCCCTAACACTTAGCACTCATCGTTACGGCGTGGACTAC
CAGGGTATCTAATCCTGTTTGATCCCCACGCTTTCGCACATCAGCGTCAGT

15 TACAGACCAGAAAGTCGCCTTCGCCACTGGTGTTCCTCCATATCTCTGCGC
ATTTCACCGCTACACATGGAATTCCACTTTCCTCTTCTGCACTCAAGTTTT
CCAGTTTCCAATGACCCTCCACGGTTGAGCCGTGGGCTTTCACATCAGACT
TAAAAAACCGCCTACGCGCGCTTTACGCCCAATAATTCCGGATAACGCTTG
CCACCTACGTATTACCGCGGCTGCTGGCACGTAGTTAGCCGTGGCTTTCTG

20 ATTAGGTACCGTCAAGACGTGCATAGTTACTTACACATTTGTTCTTCCCTA
ATAACAGAGTTTTACGATCCGAAGACCTTCATCACTCACGCGGCGTTGCTC
CGTCAGGCTTTCGCCCATTGCGGAAGATTCCCTACTGCTGCCTCCCGTAGG
AGTCTGGACCGTGTCTCAGTTCCAGTGTGGCCGATCACCCTCTCAGGTCGG
CTACGCATCGTCGCCTTGGTAAGCCGTTACCTTACCAACTAGCTAATGCGG

25 CGCGGATCCATCTATAAGTGACAGCAAAACCGTCTTTCACTATTGAACCAT
GCGGTTCAATATATTATCCGGTATTAGCTCCGGTTTCCCGAAGTTATCCCA
GTCTTATAGGTAGGTTATCCACGTGTTACTCACCCGTCCGCCGCTAACGTC
AGAGGAGCAAGCTCCTCGTCTGTTCGCTCGACTTGCATGTATTAGGCACGC
CGCCAGCGTTCATCCTGAGCCAGGATCAAACTCTCCATAAAAAATTATGAT

30 GTTTGATTAGCTCATAAATACTAAATAGTTTGTAACGTTTCGTTACTGTTT
ATTGGAATTAACGTTGACATATTGTCATTCAGTTTTCAATGTTCATTTAGT
AAAATCAATCTTTTTTATTGTACTATATTGTTTTAAGAAAAGTCAATAACT
AATTGATATTACTTAATTAATTTTAAAAGCTTTTAACGATAAGTCTTTTTC
TATATTCTTTATAATCATCCATTAAAGTTAATTCATAACCTACCTCATCAA

35 ATACTTCAGGTTCAATCTTTTTGATTTTATATAATCCCCCAAAAATAAAAT
ATTCTGG

Sequence 3357

40 step.1001h02.cons.ok
TATAACTTCCGTGAAACTTTAGCAGTTTTGCATCATCTTGTGTAATGGATC
CAGTGATTGGATTAGCTAAACTTTGTTCAATAGTTCCACGCAAAAAGTCAC
TATTTGCTTTTAAAAATTCCATTTCATCAAGATTTTTATCTAATTCTTCCG
AAATATGATTATTTGTATTAACCATGTATCCCCCGTCCCTTCTATTAATAC

45 ACGTCTCTTTGATATCTTTTATCTCTTTTCATTTGTTTTAAGTATTCTTCT
GCATCTGTTTCAGATAGGTTTTGCTCTTTGATTAACACATTTTTAATCGCT
TGATGAACATCCTTTGCCATTTTACTTTCATCACCACATACATAAATAGTA
GCGCCATTTTCAATCCATCGATTAAATTGTTCACTATTTTCTGCAATTTTA
TGTTGCACATACACTTTTTTATCAGTATCTCTAGAAAAAGCAACATCTAAT

50 TTTGATAAAGTTCCATCTTCAAGCCATTCTTGCCATTCCGTTTGATACAGA
AAATCTGTAGTGAAGTGTTGATCTCCAAAGAATAACCATGTATTTCCTTCA
AAACCTAGTTCCTCTCGTTCTTGCATATAGGATCTAAACGGTGCAACACCT
GTCCCAGGACCTATCATAATCACAGGTGTTGATTCATCTTGCGGAAACTTA
AAATTCGGATTTCGTTTTAAATAGATAGGAATTGTATCGCCCTCTTGTATT

55 CTCTCTGCAAATTGTACTGAACAAACACCTGAACGTTCCCGACCGTGTGCT
TGATATCTAACTGCTCCAACAGTAATGTGAACTTCATCGGTGTTGCTTTA
TAACTACTAGATATTGAGTACTCTCTAGGTGGTAACTTTCTTAATAATTGA
TGTAAATTTTCAGGTTGTAGTTCTGTCGTTGCGAAGTCATTTAATAAGTCA
ATCAAATCCCTTCCCTCAACGTAGTTTTGAATCCATTCTTTATCTTGAATT

```
TTTTCAGAAAGCTCTTCATTATCAAAAAATATCGCAGCATTTTCTATTAAC
GGTTTTGTTAATTTAGTAATTTCAAAATGCGATGTTAATGCCTCTTCAAGA
TTTAAAGTATCTCCATCCTCGTTAATTAAAACTTGTGTTTCTGGACTCCAA
CCTAATGTGCTAATAAGTAAGTCTACGATAGCTGGATCATTTTGAGGCAAG
ACAACTACACAATCTCCCGGTTCATATTCCTCACCAAAATTATCAAGTAAT
AACTCAATATGTCTCGTCTCTTTGTCTGAACCTCTACCATTCAAATTAATA
TTGGTCAATACTTCAGCATCGTAAGGATTAGACTTACTATATTTCTTTTCT
TTTGCTGATTTTATAGATTCGCTAACAACTTGTTCACTTTCAGTACCAGCC
GGTGTTGAATCAATCGTATTAATAACATTGGCCATCCACTTTTCAGCGTCT
TCTTCATAATCAACATCGCAATCAGTACGATGATAAAGTCTTTCTGCCCCT
AATTCAGCCAAACGGTTATCAAAATCTTTACCAGTTTGACAGAAGAATTCA
TAGGTTTGATCTCCTAATGCTAATACTGAAAATCTCACCCCATCCAATTTT
GGCGCTTTACGTCCGTGGATGTATTCATGTAATTCAACAGCGTTATCTGGT
GGATCTCCTTCGCCATGTGTAGATGTGATAATAAATAAATCTTCTACTTTC
TTCAAATTCTTAGGTTTAAAGTCATCCATTGATTTGAGCGTAACGTCATTT
CCAATATCAGATAAACGTTGTTCAAATATTTCTGCAAGTCCTTGCGCATTA
CCAGACTCAGATCCATATAAAATTGTTATCGATCTAGCTTCCGGCTCAACT
GAAGGTTCCTTTTCATGTAGCATCGCTTCTGTATTGTCATCTGAGTTATGT
TGTTCTACAGAATCTGTAGATGTATTTGATTGTTGATTTGCCATCAGATAG
CCACTTAACCATACTTGTTGGTTAGATGATAAAGTGTGAAGCAATTCATTG
ATTTGTTAGCTTGCCCTTCCGTAAAAGGACTATTGGTAACACTTAAGTTC
AATCTTATCCACCTCTTATACTTTGTTATTCCGATATTTGTCATTATTTAT
ATATGTCTATTTTTGAAAGTTATTAATTATATCTTCAATTCTAGTATGTAC
CATTTGTTATTTTCAGTAAATGTAATACTCAATTGATAACTAAAATCGAAA
AGACAATTCATAATATCAAAGGGCAATACAAACTTTTATTACTCAAAATTA
TCGTTACTAAATAAAGCTGTATCAGCTTTTAACAAACAACTCTTTATTTCT
TATCGGAATACTATGTTTTGAAATTAAAATTTTTTGACTATCTCTTATGGT
TTATCAGCTACATGTAATCCACATTCAGTCTTACTAGAATTGGACCAACGA
CCAGCACGTGAATCATTAGAATCAAATACGGGTGATGTACATGGAATACAA
CCAATACTTGGATAATTTTGATCATGTAATTCATTATATGGTAAATCCTTA
TCACGTATATAAGACCATACTTCTTCTTCTGTCCAATAGATTAAGGGACAC
ACTTTAATTGACTTAAATCTTTCATCTTTGTTAATGAAATTTGTATGTGCT
CGTGTTGGTGATTGTGCTCGTCTAAGACCTGATATCCAAGCTACAGCACCA
GATAATACGTCTTCTAGTGGTTTAATCTTGCGTATGTAGCAACATTGGTTA
GGATCATTCTTCCATAAAGCAGGATTATATTTCTCACCTTGTTCCTCTAAC
GTAAGTTCCGGTTTTTTTCATTTTAATGCGTAATTGCGGATATTTATCTTTC
ACCCTATCTATTAAGTCATATGTTTCTTGAAAATGTAAATCAGTGTCTAAA
AATACAATTTGTGCGTCGGGTTTAATTTGAGAAATCAAGTCAATCAAAACC
ATACTCTCAGCACCAAAACTGCAAGAATAAACAATATCATTTTCATAAGTT
TGATATGCCCATTTTAAAATCTCATAAGCACCTTTTGTCTCATCATTAATA
TCTAATTCATTTATAAATGGATCGTTTTGAAAATTGTCATAAGTAATTCTT
TCTATAGACATTAGCGGAGCCCCCTCTCACCTTATCAATCCATCCAAATCA
ACAAAACATATCAGAATTGTCAGAATAATTATGTTTAATACTTTAACTACT
TTTATAACGAAAGTCAAGGTAACTCGAGAAATTATCGAAATAAAGATAAGA
TTAATTTAAATTTTTTGTCCTATATGCTCATAAAAAATTTAAAGAGCCTAC
ACATATGATATGGCAGACTCTTTAATAATTTCTTTTACATTCGTTTATAGT
GACTTCATTCAAAATATTACAAATGGAACAAGATTATTTTATTCTTTAATC
CAAGAAGAATATAGGCATTTATTTTAATAATTGTCTTATTAATTTTTCCAT
TTGGTTAGGATTATCACAAGGTAAAAATCGTTTAACAATATTTCCCTGTTG
ATCTACTACAAATTTAGTAAAATTCCATTTAATTTGCGAACCAAATAATCC
TGGTTGTTTACATTTTAATAATGTGTACAAAGGGTGTTCATGCTCTCCATT
AACATTAATCTTAGCATGGATGGGGAAAGTAATACCAAACTTATATTTATA
TACTCGATAAATATCTTTGATTAATTCTGGCTCCTGATTATTAAAATCATT
GCAAGGGAAGCTCAAAATTTCAAGACCATACTTATGATATTTTTTATAAAG
CATTTCTAATTTATTAAATTGATCATTCAGAGTACAATTCGTTGCAGTATT
TACGATAATGAGCACTTTCCCTTTATATCTTTTTAGTAAATAGGTAGAGCC
ATCATAATTCTTTACCGCAATATCATAAATACTCATTTAGATACACCTCTA
CATCATTGCCAATTTCAATAAATCTATAACGCCACTAGAAGCTCCAGGAAA
TGCAAAAATATTTTTATTTAAGTCTTGTGCAGTAAGTTTTTGATTAATGAT
```
```
1
5
10
15
20
25
30
35
40
45
50
55
```

```
AAAGACTAGGAGATTAACCAAATCACCAGCGTCATTACCATAAATCTCTGC
TCCTACTAAACGTTTGTGACTATCTA
```

5   Sequence 3358
         step.1001h07.cons.ok

```
    TAGGCTAGTGGTACTAAAACACGGTCTTTATAATGTTTGAATCTGTTGTAA
    TAAAAACTATCTTCTCTATCTGCAAAATCTTTAGTTTCAGAGGTATCCTCC
    ATAAATGACCTAAATATAGGTAACTCTTCTTCAGATAAAAAGCGGACTTTA
10  ACTCCATTTTTCTTAACTTTTTTAGTATTACGCTTTCTTAAACCATCCATG
    TTTTTTAAAACATCATTAGCACTTTTGTTTGCTAAATTTAGAACAGAATGA
    TATCGGATTTGTAATACAGGATCAAATCCTTTGTGGAATCCTTCGTGTTTA
    TATCCTAAACTCTCTAATTCATCAAAAATCCAATCATGACCTGCATTTCCA
    GTTATTTCTCCCTCATGATTTAAATATTGATATGGAAGGTATGGGTCAACT
15  CTTAAATATAAACAATTATATTTTTTACATATTTACTCAATTCATTAAAG
    AAAAAATGTACAAGCTCTTTATTATTATAATCTATTACTGGACCGCGATTG
    GAATAAAAATATTTAAATATTTTCATTACAGGAACAGCTGTTAATAAACAA
    GCTGCAATCACTTCGTTATCATTATTTTTAATTCCAACTAAATGTGACTCG
    GTACCTTCAGCAACCTTTAATTCGTAATTACCTTCCATTTGTGTAAAATGA
20  CTATATGTCATACGATCAGTAAAGTCACTAAATTCTTTAGCTGTCAAATTC
    GTAAACTTCATCTTCATAACTCCCATTTATATCTTTATATCTATTTCTTTT
    TATTTTTAGAAAAATAGGTCGTTTTACTTCACCTTTTATATTTAACATTTG
    GTTCATTTAATACTATTTATATTTCTAATCTAAACAAGTATACAATAGTTG
    TAGTGAAAACTAAAATATTATCCTGCAAAACTTTTATATAAAATTAAAAGT
25  TAAATCTCGTTTATTCTGTTTATGTATAAAATAATATGGGCTAAACACTAA
    TGTCGCCCATATCACTAATTATTTACTATTATTTTAATTCATTTGCAAGTG
    CTTTACCAATGTCTTCACTTGATTGGGGATTTTGTCCAGTGATAAATTGAC
    CGTCTTTTTCTACGTGAGATGTAAAGTCATCTTTCACTACAAAATTTGCAC
    CTTGCTCTTCTAATTTAGATTGAGTTAAAAATGGTACTTTATTTTCAAATC
30  CCATTGCTTTTTCTTCACTATCAGTAAATGAAGTTATTTTGACACCATCTA
    CTAGATAGTGATTATTTGCATCTTTTACACCTACAAACGCACTAGGTCCAT
    GACATACTGAAGAGATTATTTTATTACTATTTTTAAATTGAAGTAAAATAT
    CAGCTAATTTCTCATTATTGGCAAAATCGTATACAGTACCATGTCCACCTG
    GTAGATAAATAGCGTCATACTCATCTGCATTTACATTTTCGATACTAGGTG
35  TATCGTTTAAGTGTGACACGAATTTAGCATACTGATTCAGTGATTCATTAG
    AAACAGAATTAGGATCAAGATTTACTTTTCCACCTTTAATAGAAATAACAT
    CAACATTGATACCTTCTTCAGTCAATATATTATATGGTGCTCCAGCTTCTT
    CTAACCATAATCCAGTTTCTGTACCGTCTGTAAATTGACTTGTACTTGTTA
    AAACAAATAAAACTTTTTTACTCATGAATATATACACCTCTTCAAATTAAT
40  TGTGATTATTATTTACCCACATTAGTAGTAGATTATTCAATAGTGATCATT
    TAATTTTTTGAAGTTACTATTAGTATAGTAGTTAATTCATAGAGTAATCAT
    TGTTTTTGATATGTATATTTTTCATCTATTTTCTTAACTTTGATTTAACTT
    AATATTAATTTTTATAAAGAATTCAACGTAGTTCTTATTGATTTGATATAA
    GTGATAAATTCTTTTCTTGAATCTATTTCAATACGTTTACAATTTCACTA
45  CCAATTACAATACCATCTGCAACGGACGCTATATCTTTAACATGTTCAGGA
    TTTTTGATACCAAATCCAGCAACCACAGGAATTTTTGAAACTTTTTTTATA
    TATTCAATTTTTCTCTTTAAATCTGGATGGAACTCCCCACTGTTACCTGTT
    GTGGCATTCATTGTTACCGTGTAAATAAATCCTTCTGAGTTCTTTGCAATT
    TGCATAATCCTAGCATCACTTGCGGTCATGGCAATTAACGATATTATTTTA
50  ACAGAATGATGATAAAAATCTTTTTTAAACTTTTTTGTAAGTTCGTAAGGT
    AAATCTGGAATAATTAAACCATAAACACCAGCTTCATCACACTTATCCAAA
    AATAATTCTTCTCCATAAGCACTTAGAATATTATAATAAGTCATTAATACA
    TACTTAGATGAAATAGTATTTTTATTTTTTATTAATTCATCAAAAATGAAT
    TTAATGTTTGAACCCTCGTCAATAGCGTTGCGCCCTGCTTTCATGATTATA
55  GGTCCATCTGCAACAGGATCAGAAAATGGCACACCAATTTCAACAATGTCT
    GCTCCATTCTCAGTTAATGTTTTTAAATGATGAATAAAATTTAAATCACCC
    ATAATATACGGAATGAATAATTTACTCATTTTGTTCACCACCGTTTTCTTT
    GTATTGTTTAATTGTTTCCATATCTTTATCTCCACGACCTGAAATAGTCAC
    AACAATAATTTCTTTTTCATCCATATTTGGCGCTAATTTTTCAACATAACT
```

CAATGCATGTGCACTTTCAATTGCTGGAATGATACCTTCAACTTTTGAGAA
TGTTATAAGTGCTTCCATAGCTTCATTATCTGTAGCACTTACATATGATAC
ACGACCAATATCATTATAATACGAATGTTCAGGTCCAATACCTGGATAATC
TAGTCCCGCTGATATAGAGTGTGCCAATTCAATTTGTCCATCATCATTTTG

5   AATAAGGTACATTTTGGAACCATGTAATACACCTGGTTTACCTTTCCCTAT
AGCTAAAGCATGATTATGCGTATGACTTCCTTTTCCCGCAGCTTCTACCCC
ATATAATTTAACATCATCTTGTATAAATGGATAGAACGTACCTATTGAATT
GGATCCTCCACCAACACACGCGACTAACGCATCTGGAAGTCGTCCTTCTTT
ACTT

10

Sequence 3359
step.1001h09.cons.ok
CACAACTGAAGCTTGTAAACCTGGCTTAATAATTGTATTTGCTCCTACTTG

15   TTGATCGTATTGTTCATATAAATGATGTTTAGATGCTATTGTTGGGTGTTG
TAATAACTTGATAAATGTATCATGTACATCTATATTGCTGTAATCGTTTTT
TGATGTATTATATTCCTTTTCTTCTCCTTCTAAAATATACACTGGTGCCTC
ATCAGCAAGAGGTTGCACAGGAATGTCTGCATACACTTCATCATCGTATGT
AAGCACGAATCTATCAGTGTCTGTTACTTCACCGATTACTGCACTATCTAG

20   TTCGTGTTTATTGAATAAATCTAAAAACTTTTGTTCTGTACCTTTTTCTAC
AACTCTTTAAGCCCATGTTTTTTATAGGTTCTGTTGCAAAGTAAAAAAATA
TAGCTAACCACTAATTTATCATGTCAGTGTTCGCTTAACTTGCTAGCATGA
TGCTAATTTCGTGGCATGGCGAAAATCCGTAGATCTGAAGAGACCTGCGGT
TCTTTTTATATAGAGCGTAAATACATTCAATACCTTTTAAAGTATTCTTTG

25   CTGTATTGATATTTTGATACCTTGTCTTTCTTATTTTAATATGACAGTGAT
CTTGCTCAATGAGGTTATTCAGATATTTCGATGTACAATAACAGTCAGGTT
TAAGTTTAAAAGCTTTAATTACTTTAGCCATTGCTACCTTCGTTGAAGGTG
CCTGATCTGTAATTACCTTTTGAGGTTTACCAAATTGTTTAATGAGACGTT
TGATAAACGCATATGCTGAATGATTATCTCGTTGCTTACGCAACCAAATAT

30   CTAATGTATATCCCTCTGCATCAATGGCACGATATAAATAGCTCCATTTTC
CTTTTATTTTGATGTACGTCTCATCAATACGCCATTTGTAATTAGCTTTTT
ATGCTTTTTCTTCCAAATTTGATATAAAATTGGGGCATATTCTTGAACCCA
ACGGTAGACCGTTGAATGATGAACGTTTACACCACGTTCCCTTAATATTTC
AGATATATCACGATAACTCAATGCATATCTTAGATAGTAGCCAACGGCTAC

35   AGTGATAACATCCTTGTTAAATTGTTTATATCTGAAATAGTTCATACAGAA
GACTCCTTTTTGTTAAAATTATACTATAGATTCAACTTTGCAACAGAACCT
ATATTTTGCTATAATTATTAAATTAATTGCCTTTGTGTTAGTATTCCCTGG
ATTACTAACACTTTGGTTAGCTGTTCTAAGTGATACAGGTGCAGCAATATT
AGTTATATTAAATTCATTACGTTTATTAAAAAAATAAAAAAGCCTAAATAC

40   TCTCTCGATTTGTGAGAAAGTATTTAGGCTTTTTGAATTAATAAATTTCTT
ACTTCTTATGTTTATTTAACAATTAGTACTGGGACAGAACTTCTTTTAGCA
ATCTTATGTGTAACATCTCCTAAAACATGTTTAATATTCAAATCGACTCTA
TGATTACTCATAATTACTAAATCAAATTCACCATTATTTATATCTCCAGAA
GTAATTTTTTCTAAGATTGTATCTTTAATATGCCCAAAATCTACATCTATT

45   TTATATTGAATATCTCTTTTTTCTAATTCACGTAAAAATGGAGTGAGTTTT
TTCTCTTTTTCTCTAATAATATCTTCTTTATGCTTATTATAATACTTGACA
CTAATTGCAAGATCATTTTCTGTTACAACGTGATAAATTACAACAACTGAA
TCTTCATCGGTGACTTTTTCTAAATAGTCAGGGATAGCACTAAAATCATTT
TCGAAATCATAGGGTAATAATATATTTTTGAACATCTACACCAGCCTCCTA

50   TTACTTATTATTTAGATGTCTATTACTTATATATACCCTAAATATGTATGC
TTATTCTCATATTCAAATTATTCAAATATATTCCTTATCATGTATTAGATT
CTATTTTTTATATTATTTTTGTAATTCTTTGTATAAACCTATTTATTTATA
AAGCGCTCTAATACATACCTAAATATCTTCAACTTATAAACTACTTTTTAT
TTCAATGCCTTTTTCCAAACTTCTTTTATGAGATATTGTGGTTATTACCTA

55   ACTCTCATTATTAATGTTATTATTTTATTTATACATAGTAAATGTAACTGT
AAGACATTCTCACCTAACATGTTTTGTTATTACCTGACATTTACAATTTAC
ACATTTCGCTATTATTACGCAAATTATTTTGATTTTATAATTATGATGACT
ATAATCTATTCTAATTAATCTCTCCTTTTAGGCATAAAATAATCTACTTAT
AATTAAAGCATCCACTATTTCAAAATTATTTGTAAATTGAGATTTATAAGT

GATTTTTATACAGTTATTTAAAATCTATTTCATAATCAAAAATAATAAAGG
GGGTAATAAAAATGAAAAAAGAAAAACGTTTAAATCTTATCTTAACTGTTA
TACAACAGAATCAGTTTAATAAAAAACAACAAATAGTGGATTACATGGCAA
GACATTTTGGAGTTTACTATAGCTTGACAACTATTTCTCGTGACTTACAAG
5   AATTAGAAATTTACAAAATCCCTGTTGAAAATAAAAAGTATATTTACAAGA
AAATAAATCAAACAAATCAATTAAGTGCAAAAAAACAATTAGAAATATTTA
GTGATGAGATTATTGAATTTATAACGCTAAATAACTATGTCTTAATAAAAA
CATCTCCTGGCTTTGCTCAAAGTATAAGTTATTACATAGATCAATTACAAA
TGAAAGAAATATTAGGAATTATTGGAGGTAACGATACTTTGATGATTTTGA
10   CTTCTTCAAATGAAATAGCAGAATTTGTTTGTTATCAATTATTCCCTTAAA
AATTAAATGATTATATTAATAGGAATGATTACTAATACAAAAAAATGAGTA
AAAAATCTTTTTGCTTGAATAAGATAAGGTTTTAACAATGAGATTTT

15   Sequence 3360
step.1002a01.cons.ok
AACACAACTCTTATTAACTTTTGATAAGATTTTATAGTATTAAAGTTCCAT
TCTTTACGAGTATTCATAGCTTCTAAATAATACCCAGTTATAATATCAAAT
GCTTTCTTCTCTAAGCTCTCTTCTCTATTTGAAACATGATAAGGAGCACCA
20   CCTTTAGTAACCCACTTATCTTTTTTATAATTAGATAAATAGGCAGTATAT
AAAGAGATAATGGTCTCTTTGAAACCGAAAAATTTATACTCTTCTTTTAAA
GAATCTTCACTAGCCACTATAGAAAATAGATGAGCGTTAAAAACTGAGGGG
TTATAACTTTCTGATTTAGGTGGAGCACCACCACCACCAGAATTATCAATA
GGCAAACCATCTTTATATTCTACTCCAAATAATTCTCCGATTTTAGCGCTC
25   ATCTCATCAGCATCAAATTCTATACTTTTATCATTTTCTTTTTCAAACTGT
CTTAAATACTCTTCCTGATCTTTTTCAATACCTTTACGTGTTACTTTTATC
GATCCACTAGGGAATTTGAATTCACCCACTGTAAAATCAGTCAAAATTTTT
TTTATTTGTTCTTTAAATAAAAAAATGATGATAAGAAATAAAACTGGAAAA
GAAAAATTCTGATACAAAAATTTAACGAAACTAAAAATTGATTTTAATATC
30   TCATAAATACACTCCAAAACTTACCTCCAAAATTTTAATTGACTTTTAGAT
TATTACACTAAGAAATCATTAATAGTAAAAACATATATTTATTCTAATTAC
TTTATTGACGTTGATCCTCTGAACCCTTAACAAATCCAAAACTTGTCGAAT
GGTTGGTTTAATAACTTACGCTATGCCGACATTCGTTTCAATCTGTTTGAT
AAGTTAAAATTTCATATATAGTCGTGAATTTCAATGTTCTCTTTAAAACAA
35   GGATCGTGAAAATATGCCTAAAAATCAAAACTAAATGATTCATTAGTCTTT
AATTATTTTCTGTTACTTATAGGTGTGATAACCTTTTGTGCTTATTATTTT
TTAGGTGAAGTTAATGTTAATATAGATTTATAACGTTTAAATTGCATGTTT
TAAATTGTATTTTCATGTTTGCATCTCTCTTTTTTTATTTATCTCTTAGTC
AAGATTTAAAATCTTAACTAAGAGATTTTTGTATTCTTTACATAATAAAGT
40   CTTATGGGTAGTTGAATTGTATACTCTATTCATTTACTGAATTATATCAAT
AATTTTATTCATCTTATCTTTATCACGTGATACTACGCCTACAACTCTTTT
ATAACTTTGTAATTTTTTTTGATATAAGTTCATACTTTCAAGTCTTGTAGT
CAAAGATTGAGGAGTCAAAGTAATACCTTTATTTGCTTTGACATTAGCTAA
TATAGTTTCATAATAATCAACATAATCAATATGTAACTTATCTTTCAAGGA
45   ATAATTTTTAAGAAGATCATATGTATCACATGGCGGTGTTTGTATATAAAT
ATTCTCATTTTTTAAAGATTTTATACTTACTTGGTTAGTGTTTTTAAACTT
ATTTTCATCACCATAAGCCACAATAAAGTCCTCAGTGTAAATTTCTTGTGT
ATAGAGCGAGTTATTTTTTAAATGTGTAATATCACCAATAACAACATCTAT
ATTATCTTTATAAATTTCCTCAAGTAAAATTGAAGTACTATTTTCAATAAC
50   TAAAACGACACTTTCATTAATATAATTCTTTTCATGTAGCTTATATAAAGA
AAAGCTAGGTATGACTCCTAATTTAATTCTTTTATCTTTAGACTCACTTAT
TTTACTTTTTAATTCTTCAAATTGCTCGTTTATTTTTAAAAACGAATGATA
AACAGCATTTCCTTTTTGAGTTAAAAAAATACCCTTAGAAGTTCTATAAAA
TAACTGACAATCAAAATAATCTTCAATACTTCTAATTCTTTTTACTTAAGGC
55   TGGTGTACTTATATTCAATTCTCTAGAAGCTTTATTTAAGCTATTGTACTC
GACAACTTTAATGAAGTCCTTCATATAATTAAAATTCATAAAATCCCCCAT
GGTATTAACTATTTGTTTATACCCATATTAATACATTACTTATTCATATAT
TTCAAGTAAAGGTTTATCCTTCTATTGTTCCCAAAAATAGGCAACTAATCT
CTTAATGAATCAAAATTGACTGATAGGATGAGATAATAATGAAAAAAATAT

GGATTTTAACTATAGGTATGTTTGCCTTAGGTATGGATGCTTATATAGTAG
CAGGATTAATACCTTCAATAAGTAAAAGTTTTAATAAAAGTAGCTCTGCTA
TTGGGCAAGGAGTAACAGTTTTTACATTGTTTTTCTCTATCTCTGCCCCCA
TTTTTTCAACAATATTAGCTAAATCCCCAGTTAAAAAAATACTAATAATAG
5 CATTCAGTATATTTACTTTAGCCAATATTATAACCGCAATATCTATGAACT
ACATGCTATATATCGTATCAAGAGCAATCGCTGGTTTAGGAGCTGGCGTAT
TCTCACCAATTGCAATAAGTGCAAGCAATCATTTAGTCTCCGAAAAGCATA
AAGGAAAAGCAATCGCTTTTACAGTAGGCGGAATGAGTGTAGGAACTGTTA
TAGGAGTTCCTCTCGGACTAGAAATTGCCAACATTTCTAATTGGCGATTTG
10 CAATGTTGGTTATTATTGTCATTAGTTTTATTGCATTAATAAGCATATCTA
TATTGATGCCTAAATTTAAAATAGAAGCTCCTCCAAATTTAAAAGATCGTT
TTCAATTATTTTTAAACAAGCATGTACTAAGAGTTATTTCGGTTACATTAT
GCGCTGCCATTGCTAGTTTAGGTTTGTATACTTATTTAGCCGATATTATTA
AAACAAATACAGATACAAAAAATTTAACTCATTACCTTACAGCGTGGGGAA
15 TAGGCGGATTAATAGGAAGTTTTGGTATAGGATTTATTATAGATAGATTTA
AAAATACAAGATTTGTTATGCTAATTATTTTAATTTTACTAGCATTAAGTT
TTGGTTTAATTCCTATTTCTATTAACTTGCCTATATTAGGTTTAATTCCCT
TTATTTTATGGGGAGCTATGGGATGGGCTACACAAGCTCCTCAACAACATA
TATTATTGAAAAAACATCCTGAATATGGAGGCTCTGCTGTCGCTTTAAATA
20 GTT

Sequence 3361
step.1002a02.cons.ok
25 TTAATTAATTTTATTTATTTAATTAATATTTTGTAAAATTAAATACAATAA
TACATATTTATGATATGTTTTTAAATGTAATAAAGTCAAATAAAAGGAGGA
AATTTATGAAAAAGTTAATGATGAGCTTATTATGTTGCACAATATGTGGAA
CAGTTTTAACTACAGGAAGTGTAAATGCACAAGGTGATAGCTCAACTAACT
CTGAGTCGTTAAAGGAACTTCAAAATGAAGGTATTGTGTCTAAATCAATTA
30 CTGAACAACAATGGCAACAAATGAAAGCTCAAGAGCGTAAAGATGAAGCAG
AATTTGAGAAAACAGCTGAAGTACAATGGCAGAAACAACAAAAACAAGATC
GAATAGATCGAGAAAATCGTTCAAGAAAAAAGAAATTTCATTTGAAAAAAG
GTGATTTTTTCATCACAAATAACGTAAGTTCAAAAGGGTTTACAGGTCATG
CTGCAATATATACTGGAAAAGGAAAAGTTAAAGAAGCGCCTGGATATGGAC
35 AACCTGTGAGGGTAAAAAGTTTTAGTGATTGGAAAAAGAGTACTTTGAAAA
AAAGAAAAGGAAGTCCCAAACATCGTTATATCAAGGTTTATCGAGCACCAA
AAAAATATAGAGGGTAAAGCTGGAAACTATGCGAAATCTCATTTTAACGGTG
TACCTTACAGTATAACGACGAATCCATATTCTAAAAGCGTTACATACTGTT
CGAAACTTGTTTGGCAATCGTATTATTATGGTGCCGGACGATTATCAGTTC
40 TTCCGGTTGTTACATCTCAATTTATTATTGAGCCATATAGTCTCAATAAAT
ATATTCCATCAAAAGCAGTTAGGTCTTATAAAAGAAGCTAAAAAATTAAGG
TGTTAAATATATAATTTATACATTTTAATTACCCATAATCGTTAATTATGT
TACTTTTCTTTGATTTCATCGTTTATTAGAACACTTTGATAATACATCATT
TATGAACTTCAGTACATATTATGTGATAATCGACCATTAAAAAAATACGATT
45 AAATGATTAAAAGAACAATAAAAAAGCAAGCGACTAATAATATGATTATTA
GTCGCTTTCAAACTGTCAATAAGCGCTCTTTTCAGATACTTTATTTATCTA
ACTCTACTGTTTTTATCTTTGGTATTTCTGGAAGAAATGCATACTTAATCT
TTATAAATAATGAAGCTGGCTTTCCATTGATATCAATACGATAATAACCTA
TATGCCCACTTCCTTGATTATCAGATTTTCCATCTATAGAAATTTTATTTT
50 GCTTTTTTAAGAATTGATACGTTTTATTATCTTTAGCAACCTCTTTTAATT
TCTTAGTATCTTCATGCTTAACATAATTATTTAATTGCTTTTCACTTGAAT
GACACGCCCATGAATCAAAATAGAAAAAGTTTACAAAAAATATAACAAATTG
CTAAAAAAATTATTATACCAATAACATAGAATACTTTTTTTCATTAAAGTTC
ACCTAGTTTTCAAAAAAATTTTTATAATTTATGATAACTTAAAAGTGTATA
55 TTTTTTCTACATAATATATTTAATTTTCAAGGAGATGTAAAAAAGTTGAAA
AATTTCGCAAAACTAATTTTGGTGGGCATTTTAGTATCGGGGTCAGGGATA
GCGAGTGTACAAACAAATATAACTCACGCAAAAGAAAGTCACGATTCAACT
CCTCAAAATATTAAATTAGTGGGAACGTATGATACTTCTCAAGTTGATTCC
AAAACGATGAAACAATTTAAAGAAATAGAAAAAGAAGATAATAATTTCCAC

ATAACTAAACATGGAAATAAAGTCGTTGTAGAAGACAAATTACCTAATCCA
GAGAATAAAACTTCAAGTTATTCAGCTGATGGTAGTGCTGAAAATAATACA
AAAGTAATTAATTTCTCTGATTTTGTTGGTAATATGGATGGGAAAGATGAT
GGAAAAATATCGGATGGGATAACCTTTTATAGTGGTAAATCATATAACGGA
CAACACGATGGTCAAAAAGTAAAAAAAGGGACTCATGTACATTGTAATAGA
TTTAACGGAACAAAATCTGATCATAGATACTGGTCAAAAAAACATCCTAGA
GCTTATGTAGATTTTTATAAAAGTGATTGCTGGTATCACGCCAAAGCTTAT
AAATGTTCTTCCTTGGGAAAAATGACTAAATGCGATGGTTTGAATAGTATT
TATAGAAAAGGTGTCAAAGATTGCTCATCATGGAAAGGTAAACCCAAACAT
AAAAACTGGCCTAAAACAGCATGGTATAGAAATTAAAAAATGAATAAAATC
TTAAAAATATTAATAACTTCTATTATTGTTATCATTATTACCTTAACAGTT
TGGACTTTTAGTGTGATTACTTATCAGAAACACAAGAGTGAGAAAATCATC
AATCACGTTATAGAACGTAAGGGTTGGATAAAAAAATAAAAAATGAAAAA
ATGAGTTTTAATATTATAATGGGATATGCTGAAAAGATATTGTTTTTAAA
GATCAACCATATAGTGAGTATGAGTATAACGTGACACCAGCACCATGGACA
GATGATAAAGAATATAAGGTGTGGGGGGAAACAGATTTACAAAAGAAAGAC
TCCTATTATAAATATCTTTTAGAATCAGAACCTTACAGAAAATAATATAAA
TGTATAACATAATAAAAAATAGAGATCATTCATGGACTTCTAGTAAGATTC
AAGGTAGAAATACAGATGGTGGATTAGAATGGTCACCAGATCATAAATCAC
TTATTTATAAATATGATGCAACATTAGGTAGACAAATAAATACTAATGACG
TGTTAACTTTACTTCAAGCAACAGCTAAAAACTCAAATTTACGTTCAAATA
TCAATAGTAATGAAAAACAGTTAGCAGAACGAGGGTCTAATGGGTATTCTA
AATCTATAATTAGAGATGATGGCGAGAAATCTTATTTACTTAACTCAAATC
CTATTCAAGTATTAGACTTAGTAGAACCAGATAATGGTTACGGTGGACGTC
AAGTCAGTCATTCTAACGTTATATATAATGAAAAAAATTCTTCTATCGTAA
ATGGTCAAGTTCCAGAAGCTAATGGGGCATCCGCTTTTAATATTGATAAAG
TTGTTAAAGCTAATGCGGCAAATAATGGTATTATGGGTGTTATCTATAAGG
CACAATTATACTTAGCACCATACAGTCCAAAAGGTTACATTGAAAAAATTAG
GCCAAAATTTAAGCAATACCAATAACGTGATTAATGTTTATTTTGTGCCTT
CTGATAAAGTAAATCCTAGTATAACTGTAGGTAATTACGACCATCATACGG
TATATTCTGGTGAAACATTTAAAAAATACTATCAATGTAAATGATAATTATG
GATTAAATACAGTAGCTTCTACAAGTGATAGTGCAATTACTATGACCAGAA
ACAACAACGAGTTAGTAGGTCAGGCTCCTAATGTTACTAATAGCACAAATA
AAATTGTAAAAGTTAAAGCCACAGATAAAAGTGGAAATGAAAGTATTGTTT
CTTTCACAGTAAATATAAAACCATTAAACGAGAAATATAGAATAACAACTT
CATCAAGTAATCAAACACCAGTGAGAATTAGTAATATTCAAAACAATGCTA
ACCTTTCAATTGAAGATCAAAAATAGAGTAAAATCTTCACTCAGCATGACTA
AAATTTTAGGTACAAGAAATTATGTCAATGAGTCAAATAATGACGTTCGTA
GTCAGGTTGTAAGTAAAGTAAATAGAAGTGGGAACAATGCTACAGTTAATG
TTACAACTACATTTTCTGATGGTACAACTAATACAATAACCGTTCCAGTTA
AACATGTGTTATTAGAAGTTGTACCTACTACTAGAACAACAGTAAGAGGAC
AACAATTTCCAACCGGCAAAGGAACTTCCCCAAATGATTTCTTTAGTTTAA
GAACGGGAGGTCCAGTTGATGCGAGAATAGTTTGGGTTAATAATCAGGGAC
CCGATATAAATAGTAATCAAATTGGTAGAGATTTAACATTACACGCTGAAA
TATTCTTTGA

Sequence 3362
step.1002a06.cons.ok
GGTGCGATATATACTTTATTTCAACCATATTTAAATACTATGATGCACTAT
TGGTTAATCAAATAAGATTATTATGAGAAAAATAACAATTAAAATTACGTA
TACATAAGGAGCGTTATCATGTTAGACAAAAATCAATTGGAAAAGTATAAC
CAAGAGCATTTGTATGAATATGAAAAAATTAATGAGTAGTAATGAAAAGAAT
GCTTTAGATGAAAAAGTAGATCAGTTAAATCTTGCAGAAATTCAAGATTTA
TATCAAGATTTATATGTTAATAGAAAAACTATTGATGATGTATCTTCTGTA
TCTGAAGTCAAATATGAAGTGAAATCACGACTCAATGAAGAAGAACGACAT
ACATATGAACAAAAAGGTTATGGGCAATACGAAATGGTGAATTTGCTGTA
TTATTGATGGCTGGAGGACAAGGTACGCGTTTAGGATATAAAGGGCCTAAA
GGTTCTTTTGAAATAGAGGGTACGAGTTTATTTGAACTTCAGGCGCGTCAA

CTGATTCGTTTAAAAGAAGAAACCGGCCACACAATTAATTGGTATATTATG
ACAAGTGACATTAATCATAAAGATACAATAGAGTATTTTAAACAACATAAA
TATTTTAACTATGATGCCAATCATATTCATTTCTTTAAGCAAGATAACATT
GTTGCTTTAAGTGAAGAAGGAAAGCTTGTTTTAAATAGAGATGGACATATA
5 ATGGAAACACCTAATGGTAATGGGGGTGTATTCAAGTCTCTTAAGAAAGCA
GGATACCTTGATAAGATGCAACAAGATCACGTCAAATATATCTTCTTAAAT
AACATTGATAATGTCTTAGTTAAAGTTTTAGACCCGTTATTTGCCGGTTTT
ACAGTGACACAAAGTAAAGACATCACATCAAAAACAATTCAACCTAAAGAT
AGTGAAAGTGTAGGTCGGCTTGTAAATGTTGATTGTAAAGACACTGTGTTA
10 GAGTATTCTGAATTAATGTCATATTTCCAGGAGCTGCACTCCAGTTCATTT
CTGAAACTAAAATACTTCCATCGCTATTCACGCGCTCTACAAACGCTACGT
GACCATAGTAACCAGCGTCAGTTTGTGCAATTGAGCCTACTGTAGGACGAT
AATCAATAGTATATCCATCAGCAGCTGATGCATTGTCCCAATTA

15

Sequence 3363
step.1002a08.cons.ok
TTACCTCCATTTAAGATTCAAAGTGATATGTCCAAAATGGCTCTATTTTTA
AGAAACTTTGTGCTTGACCATCTATTTCGGGACTAGCAAGTTGTTTTAAAA
20 ATGGTCCTGTTTGTGAAGCATGTGCTTCAAAAGCTTTAAGTTTTATATCAC
TATATTGAGATATGTCATTTTGAATATCAGGTTCTCCAAGAATTTCTGATG
CATCATTGCTAAACGCTACAAGTGTAAGACGGGGTCGATCTGATTCATGCA
TGCGTCCAACTGTACGTACTACAGCTTCTGCAGTTGCCTCGTGATCAGGGT
GAACTGCAAATTTAGGATAGAACGAAATAATTAATGATGGATTTGTTTCGT
25 CAATAAGTGATTGAATCATTTGATCCATTTGATCGTAAGGTTCAAATTCAA
CAGTTTTATCTCTTAACCCCATTTTCCTTAAATCTGTAATCCCAATTGCTT
TGCATGCTTCTTCTAACTCACGTTCACGTATAAATGGTAAAGATTCTCTTG
TTGCAAAAGGAGGGTTACCTAGATTACGTCCCATTTGTCCTAGGGTAAGAC
ATGCATATGTGACGGGAATACCTTTTTCAATATAACTTGCGATAGTTCCAG
30 CAGACGAAAAAGTTTCATCATCAGGATGGGGGAAAATCACAAGTACGTGTC
TTTCATCAGTCATGATTGGCAACCTCCTTTATATAGTAAATGGCTGTTTAC
TTATTTCTATTGTAGCAGCCAATTGTCCTTCATAATTAAAACCTGCAATTA
AAAATTCGTTATTCTCATTAACTTCATAGTGCGTAAGTCCTTGCACGTAAA
CCCAGCCACCATCTTTTAATTTAAGACCTACACGATATGGATCTTTGTCGC
35 CGCCTTTAAGTTGTGCATGTTCAAAAGTCACGACAATATTTCTTAAAAATG
TTCCAGCGTTAAATACGCGTTGATCGAAATGATTTGCATAAGCACCATTTG
TTGTTTCAACGTGTAGGTAAACAGGCTGATTTGAGTAAGAAGTTAATAAAT
CTAGCACTTCTTGTTCTTTAATTGGTTCCAACACGTTCACTCCCTTAAACT
AAACAATATGTCAGTGACGGACTACCGATAAACATATTATTTACCAGCAAA
40 CTCGTTTATCTATCTATTTTACTAGAAAGCTATGTATATTTGTACACGAAT
GCTCGATTTAATTAGGAAAAAATTCAAGAAATAACGTTGAATATTACATATA
GAGCGTTAATATGTTTAAAATCTTAATCATTTGTTACAGTATAATTTAAGA
GATATAAAAGAGTAAGGCGTTAGTCATCAATCATTGTGATATTGAAATTAG
AGGTGTTTGATATTGAAAATTATTAATTTAGATTCAAAAAATCTTGCGTCA
45 TTTTATGTAGCATGCGAGTTGTTTAAACAAATACAGCAGCACCCTCATGCC
AAACTCGGTTTAGCAACTGGTGGAACTATGACTGACGTATATCATTACTTA
GTAAATTTATTAATCAAAAATAAAGTTGATGTGTCGCAGGTTGAAACATTT
AATTTAAATGAATATGTAGGTTTAAAAGCAAGTCATCAACAAAGTTACCAT
ACCTACATGAATAAAGTGCTATTTGAGCAATATCCTCATTTTGTTAAAAAT
50 CACATTCATATTCCTGATGGCTTATCAGAAAATCTAGAAGCAGAAGCGGAA
CGATATAACAATTTATTAGATGAAAGAGGGCCAATCGATATTCAAATTTTA
GGAATTGGAGAAAATGGTCACATTGGTTTTAATGAACCAGGGACTGACTTC
AATAGTGAAACACATGTGGTGAACTTAACAGAAAGCACCATAAAAGCAAAT
AGTCGATTTTTTGACAATGAAAAGGATGTTCCTAGACAAGCAGTTTCAATG
55 GGGGTAAAAAGTATTTTAAAAGCAAAAAGGATTATCCTACTCGCATTTGGT
CCAAAGAAAAAAGAGGCTATAAGTAAACTGTTAAATGAACAGGTTACCGAA
GATGTACCTGCGACCATTTTACACACACACCCTAATGTTGAAGTTTATGTA
GACGATGAAGCAGCGCCAGATTGTTTATAATTATATAGTTGTTTTATATAC
AAAATATTTTATATGTTTAATAAGTTTAGCGAAGGGTATACAACTATAAGC

AAAAGAGAATTTCCCAAAACGAAAGGATGAATTTCATTGGAACTACAATTA
GCCATTGATTTATTAAATAAAGAAGAAGCAGCAAAATTAGCTCAAAAAGTT
GAAGAATATGTAGATATTGTTGAAATTGGTACGCCAATTGTAATTAATGAA
GGGTTACCTGCAGTTCAACATTTAAATGAAAATATTAATAATGCTAAAGTA
5 TTAGCTGACTTGAAAATTATGGATGCAGCAGATTACGAAGTGAGCCAAGCA
GTAAAATATGGTGCAGATATTGTTACAATTTTAGGTGTTGCTGAAGATGCT
TCAATTAAAGCAGCAGTTGAAGAAGCGCATAAACATGGAAAAGCATTGCTT
GTTGATATGATAGCAGTGCAAAACTTAGAACAACGTGCTAAAGAACTAGAT
GAGATGGGTGCAGACTATATCGCAGTTCATACAGGTTACGACTTACAAGCT
10 GAAGGAAAATCTCCATTAGACAGCTTGCGTACAGTTAAATCTGTTATCAAA
AACTCTAAGGTTGCAGTAGCAGGTGGTATTAAACCAGATACTATCAAAGAT
ATTGTTGCTGAAGATCCAGATTTAGTTATTGTTGGTGGCGGTATTGCGAAT
GCTGACGATCCTGTAGAAGCAGCAAAACAATGTAGAGCAGCTATTGAAGGT
AAATAAGATGAGTGAATTTAATAATTATCGTCTTATTCTTGAAGAGTTAGA
15 TTCTACTTTATCTCAAGTAGATAATACAGAGTATGAACGTTTTGCTAATGA
TGTTATAGGTGCAGATCGCATATTTACAGCTGGTAAAGGTCGTTCAGGTTT
TGTTGCTAATAGTTTTGCAATGCGCTTAAATCAATTAGGTAAAAATGCCTA
CGTTGTAGGTGAGTCAACAACACCTTCAATTAAAGAACATGATTTGTTTAT
TATTATTTCAGGTTCAGGTTCTACAGAACATTTAAGATTATTAGCTGAAAA
20 AGCACAATCTGAGGGTGCAAAAAATTGTCTTATTAACTACAAATGCGGAATC
GCCAATCGGTAATCTTGCAGAGACGGTTGTTGAATTGCCTGCAGGTACTAA
ACATGATGTTGAGGGTTCGAAACAACCACTTGGTAGTTTATTTGAACAGGC
TTCACTTATATTCTTAGATAGTGTTGTATTACCTTTAATGGATGCATTTCA
CATTAGTGAAAAAACAATGCAAGAGAATCATGCTAATTTAGAATAACTAGA
25 ATGAGAGATGAGCACTT

Sequence 3364
step.1002a09.cons.ok
30 AGTATCAGAATTGTCATTAGATGTGTCGTCATTACCACATGCGCCTAAAAT
TAAAGCACTACTAAAAATTAAGACTAAAAACTTTTTCATATGTAATTCTCC
TTTACTATATATCTTTATATTCAAACACTCGTAACGGCTCAAATTGAATAA
CGTATTTGCTATACCGAGTGGAATAACCATATTTTTGTTTATAATGTTCAA
TACTTTGTAGGACAAAACCTTCTGTAACCTCAAAAAAATTAGCAAGTTCAT
35 ACAAGTCATGAATGCCTTGCAAAAATGCTTTTACAATATCTTTAAGAGGGA
TTAACTTTTCATAGGCTAACCTACGTGCATAACCTTCAAATTTTCTATGAT
TAAAACTACTTTGATCAACTATATTTCCATATGTAAGTTCGTTATGAGCGA
GTTCCTCTGCAAGTGTTTCAAGTTTGGAAGTAATAGGCAAGTTACGATTAA
TTAAAATCATATCTCCAAGCCACAAACCAGATAACCTTTTAGGTAAGTTAT
40 CACATTCAATGACTTCAATATAGTCATGTTCAATTAACATATCTTCATATT
TCCCCACATAAAACACCCTTTATTTTCTTTTACTTCTTATGTATTCAGCGT
AATCAAGAATTTCTTGCCATTCTTCGTCAGTTAAATCACCATCAAGATGCG
CTGCACGATGTTGTGGTTGTTCCTCAACTTCTTCTATCCAACCCATCAAGA
AAGCAGGGTTAACATTTAATGCAGTAGCTATACTTTCAATAGTATCATTTT
45 TAAGATTTTTAATATTCCCGCTTTCATAACGTTGTACAGTAGCTTCAGTTT
TACCGATTTTTCTTCCTAGTTCAGCTAAAGTCATACCTTGTTTTTCTCTTG
ATTGTTTCATTCTTTTTGAAAAGCACATCGTAATACAGCTCCTTTTACTTT
ATAACTATATTATAAGGAAAACTTTCGTGATTTGCAATATTTTTTCTTAAA
AACTTTCGTAAAATGCTTGACCTATTTTGTACAGCCATGATAAGATTACTT
50 ACGTAATACGAAAGGTGGTGAAATGAAATGCCTATCGACACTAAACTTTTA
AAATCAAAAATGGCTTTAAAAGGACACAACATCAAAACTCTTTCTGAAGAA
ATAGGTGTTAATAGAGATACATTGTCCAATATGATACACGGAAGAACTAAA
CCTTCATATCCAGTAATAAATGGAATTTATTTCGCGTTAGATTTAACACCT
CAAGAAGGAAGAGATATTTTTTTTTAACAAAGACTTACGCAAAAAGAAAGTT
55 TTAACTTAAAGGAGGAATTCAAATGACTGAAAACAAGGTATTATACAGGGT
TTTTGTAAAAGAAACAAAGAATAGTGGATATACATTAGTTTTTTGAAAGTTT
AGAAGAAGTCGCAGCTAATATAAAAAAATGCAGAAGAATTTGACGAAATTAT
AATAAGCCCCTCAAAAAGAGAGGCTCAAGATTAACGATACTAAAGATAATC
TTTTAAATGTTGTGAACTGCAAATTACATTAGACCATGCGGCTTCACCAGT

CAGCTTACCAACGAACAAACTATCATCGCTATCAATAAACATTTTTAATTC
GTTTCTAATAGTTTTAGCATCATTAGTTGAATTAATAAACCAAACTGATTT
ATTGATTTTAGCAACATTAGGATAATCCTCAATACGCTCTATTAATTTTGG
ATAATCCTTTTGGTTATTTAAATCATACGCAATAATAAAACTTTTCATATT
TTCACCTCCTTTTATAAGGAGTATAGCAGAAAGGAGCATAAACAATATGCA
AGATTTACAAGTATTTAATTTTGAAGATTTACCAGTAAGAAAAATAGAAGT
AGATGGAGAACCATATTTTTTTAGGTAAAGACGTGGCAGAAATATTAGGTTA
CACAAGATCTGATAATGCAATTAGAAATCATGTTGATGATGAAGATAAGCT
GACGCACCAAGTTAGTGCATCAGGTCAAAAACGAAACATGGTAATCATCAA
CGAATCTGGTTTATACAGCTTAATCTTTGACGCTGCTAAACAAAGTAAAAA
CGAAAGTATTAGAAAGAAAGCTAAACGTTTTAAACGTTGGGTAACCGAAGA
TGTTTTACCTTCCATTCGTAAAACAGGTACTTATCAAGTTCCTGATAATCC
AATGGACGCATTGCAACTTATGTTCGACGCACAAAAACAAACCAAAGAAGA
AATAGCAACTGTTAAAGCAGATGTTATTGATATCAAAGAAAATCAAAAGCT
AGATGCAGGAGAATACGGATTGATAACAAAAACAGTTCATCAACGCGTTGC
TTATATCAGACAAATTCACGGACTACCTAATAATAAAGAAGTTAACAAACC
TTTATATAGAGATATTAACAGTAACGTAAATACGATGGCTGGTATTAAAAC
AAGAACACAATTAAAACAAAAACATTTCGATGACGTAATGAATATGATCAC
AAATTGGTTTCCATCTCAATCAACAATGTATGTCATCAAACAATTAGAAAT
GGACTTTGAAAACGAAGTATAAGGAGTGATAGCAATGGAATACATTGGATT
TGCGGACGCTATCGAGTTTGTGAAAATAAGTGGAATTTCTAAAAACGATTT
AGAAAAGCACGTTTATAGCAATAAAGAGTTCCAAGAGAAATGTATGTACAG
ATTTGGCAAGAATCATAAACGCTACATCAAGATTAGACCGGCAATTGACTT
TATAGAACAAAATTTAATGGTGTCAGAAACGTCACTTTAGAGGAGTTTTAC
TGAAGAAAGAGGATATATCGAAAAACTTTTAGAAATGAAAGGAGAGAATGA
TAAATGAAATATCTATTAAGTTACATGACGATGTTTATCGCAATGATCATC
ACATTACTTTTAGGAGGTGGTTTCTTTACGGTAATAGCATTTTCAATGTTA
ACACTTATCTTTAGCACATTCTTCTGGGAAAAGTGGCTTGAGATAACAAAA
AAGACTGAAACTTGCGCCAACAAGTAACAGTCAGAATCTAATCAAAATATA
CAACTTAATTTAATCAAAATATACGGAGGTAGTCAAGTTGAGACACAAATT
ATTAAAAACGGCCAATGATTTAAATA

Sequence 3365
step.1002a10.cons.ok
TCTATTTCTTTAAAAATGTTTGTTTCTAAAGATTTTAGCGCACGTTTTTTA
GTACGTACTTTTTCGTGTGGATGTATTTTACCATCTTCTTCAAAACGCAAG
ACAGGTTTCATTTTCAATAATGTACCTACCCAAGCTTGAGCTCCAGTGATA
CGACCACTTTTTTGTAAATTTTTTAAATCATCAACAATTAAGTATGCACCA
ATATGTTGTCTTATTTCAGTTAGTTCATTAATAATATCATCAGGTTTATAT
CCCTTTTGTACCAATTGAGCAGCGTAAATTGCAAAGCTACCTTCAATCATC
GCAGCAAGACGGCTATCAAACGTATGTACTTGAATATCTTCAACCATTTCA
CCAGCTTGTGTTGCTGAAGGATAGCTTCCGCTTATACCACTTGATAAGTTA
ATCACGATGACATCAGTGTATCCTTGTTCTCTTAATCTCTCAAAATTTTCA
ATCCAATCGCCAATAGCAGGTTGGCTTGTTGTTGGTATAGTTTTAGATGAA
GCCATTTTTTTATAAAAATCATCTACAGAAAAATCATCACTCTCAGTGAAA
TTCACTCCATCATCGAAAGTTACACTTAGTGAAGCGACTGGTATGTTATAT
TGTTCTATTATATGTTGTGGTAAATAACTTGTAGAATCGGTCATAACTGCA
ATCTTCATCTTAATACTTATCCTCCTCAATTATACTCATAGTTATGATACA
ATAATTTTTTAAAAAAGAAAATATCATCGTAATAATTTAGAACCTATCTTG
TTACGTGTATAATAATGTTAGAAAACTGAATTGTAGGTGGCAAAAATAAGA
TGGATAAATCCATAATTACTATTAAACAAGCACATTCAATTGAAAATGTGA
TAAGTAAATCACGCTTTATAGCATATATTAAGCCTGTTTCGACTGAAAATG
AAGCAAAAGCTTTTATAGATGAAATTAAAACAAAACATAAAGATGCAACTC
ATAATTGTTCAGCCTATACTGTCGGGACCAGAGATGAATATTCAAAAGGCAA
ACGACGATGGCGAACCAAGTGGAACAGCTGGCATCCCAATGCTTGAAATAC
TGAAAAAACAAGAGATACACAATGTTTGTGTCGTCGTGACACGCTACTTCG
GTGGTATCAAGTTAGGTGCAGGCGGTCTTATTAGAGCATATAGCGGCGCCG
TGCGTGATGTGATATATGATATAGGTAGAGTCGAACTAAGAGAAGCTATTC

```
      CAGTAACCGTTACGTTAGATTATGATCAGACAGGTAAATTTGAATATGAAC
      TTGCCTCTACTACATTCTTATTAAGAGAACAATTTTATACCGATAAAGTAA
      GTTATCAAATTGACGTAGTAAAAAATGAATATGATGCTTTTATAGACTTTT
      TAAATCGAATTACTTCTGGAAATTATGATTTGAAACAAGAAGACCTTAAAC
 5    TATTACCTTTTGATATTGAAACCAATTAAATAAAAAACACTATGCAAGTTA
      AAGCAACGCGTGCTTATGACCTACATAGTGTTTTTTATTATGAACGGTGA
      TGCTCATCATAATGATGTTGCTTACCATTTCCTTTTTTTGTAATTAAATTT
      AATATTGGTCGATAATTATCATCTATTAATCCAGTAAATTCAACGATCAAC
      TCAATCGTAAAGACAATGAGAATGAACATCATAAGTGCACCCAACGGTTGG
 10   GATAAATAGAGGATAACACTAGATAAACTAAACATAATCGCTATTGAATAA
      ATAAGTAAAACGGTTTGTCTATGCGTATATCCTAAAGCAAGTAATTTATGA
      TGTAAATGTGACTTGTCCGCTTGCATTATATGTTGCCCTTTTTTCATTCGA
      CGAATCATTGCAAATAATGTATCAATAAATGGCACCGCTAATATAACTATA
      GGAAAGAATAATGCAATAAATGTGATATTCTTAAAGCCGAGTAAGGATAAG
 15   AAACCGATAATAAATCCTATCATTAATGCACCACTATCACCTAGGAAAATT
      TTCGCTGGGTGAAAGTTATAGAATAAGAAACCAAGTAAAGACCCTAAAAGT
      ACACAGCAAATCATGATAATAAATATGTTCGCTTGTAAAATAGCGATGAAT
      CCAATAGTCATTAATGCCAATGCTGAGACGCCTGAGGCAAGTCCATCAAGT
      CCGTCGATAAGATTAATAGCATTGGTAATTGCTACAATCCATATTACTGTA
 20   ATAGGAATGCTGAATATGCCAAAATGAATCGTTGGACCAATTGGCAATGAA
      ATAAAGTCTATTGTAATTCCATAAAACGTAACAATTAAAGCTGCAACAATT
      TGACCTGCTAACTTTAAATAAGGTCTTAGATCGTAAATATCATCAATCAAT
      CCAACCATATACATTGTAATTGCACCTAATATAAGCGGTTTAACCTCACGT
      TCAATAGGGTGTCCGAGCCAAATCCCTATTAAGAAAGAAAATAAAATGACC
 25   GTTCCTCCCATCACTGAGATAGGTTTCGTATGTACTTTTCTGAAATTAGGA
      CGATCTACTAAATCTAATTTTTTTGATATTACAATAATAATGGGTGTAATT
      ATTAAACTGACTATCATAGTAAAAGCTATAAGTAATAGTGTATACATCAGT
      TCACCTTCATTGAATAAGTTGTACCTCACGCTACCACTGTTTTGTGATTTG
      CTTTCACTTTTAGTTCTTTCACCTGAAGATTTTTGACTATTTTCAACATCA
 30   TCTTTTCTTAAAACATCATTGCCTTTTCCGGATACTTCGCTTAAATCAACG
      CCATTTTTACGCGCATGTCGACGTGCAGATGGTGTTGCATTGATACGTTGC
      TGGCTATTATCTTGTGTATTTTCACTTTCACTAGAATTCGGATTTGATTGT
      TTTGGCGTTTCTTTCTCTTTAGCTTCATCTTTTTGTGAATTTTATTATCTT
      TAGCAACCTCTTTTAATTTCTTAGTATCTTCATGCTTAACATAATTATTTA
 35   ATTGCTTTTCACTTGAATGACACGCCCATGAATCAAAATAGAAAAAGTTTA
      CAAAAATATAACAAATTGCTAAAAAAATTATTATACCAATAACATAGAATA
      CTTTTTTTCATTAAAGTTCACCTAGTTTTCAAAAAAATTTTTATAATTTATG
      ATAACTTAAAAGTGTATATTTTTTCTACATAATATATTTAATTTTCAAGGA
      GATGTAAAAAAGTTGAAAAATTTCGCAAAACTAATTTTGGTGGGCATTTTA
 40   GTATCGGGGTCAGGGATAGCGAGTGTACAAACAAATATAACTCACGCAAAA
      GAAAGTCACGATTCAACTCCTCAAAATATTAAATTAGTGGGAACGTATGAT
      ACTTCTCAAGTTGATTCCAAAACGATGAAACAATTTAAAGAAATAGAAAAA
      GAAGATAATAATTTCCACATAACTAAAC
```

```
 45
      Sequence 3366
      step.1002a12.cons.ok
      GGGAGTGGGCATTTATGATTCAAATAAAAGGAGCAGTCGATTTCCCTATTT
      CATTGGATAGTACGACTTGGATATTTGACGATAGAAAAGTTACTATTGATG
 50   ATTTAGAACGTGGGGTATTTGATGGTACTAGACCCATCAACTTTGATGATA
      ACAAGGAATGGAACCGTGCTATTTTAGAAGGACAAACAAATCCACCAACGC
      TTGATTCTGAAATTAAATATAAAAAACGATCGGTTTTAGATGAAACATTTG
      TGATAAATATGACACCATTTTTCAAAAACGCCGAACCTATGGAAAATGCTA
      GCACAATAAAACTTTCTAATAAAGATTATTCCATCAATGTACCCATGGACT
 55   TATTACCTTATTTGTTCTTTCAATTTTCTAGTAATGGCAAACGTTTATATA
      GTGATAATGCTGTTGATAGTTTCATATATATTCCTGAAGAAGGATACTCTC
      ATCAATTTAAGTATGTCACACATATAGAGGTGATTTAATATGAGAAAGGTT
      CAATGTATTATTTGCGATACTAAAGTTTTTATTGATGAACATACAGTCGAG
      GCTAAACGCTTAAAAAAATAATCCTATACGCACTTTCATGTGCGATGATTGT
```

```
AAAAGTAGACTCGATACACCTAAACAAAGCAAACGAGATTAAAATGATATA
TTGGAAATTTCAAACTATCTCACATAAATAAAGAGTGTGGGACTTAAAGTT
CCAAGATAAATAAAAAAGACAATTTCTATTGAAATTATATAGAAATTGTCT
TTTTTATTATTTTTTTGATTATTTTCAGCTCGTTGAGCTGTTACTTTTCTT
ATATTAAGTGCCATTAGCACAAATCCTAGTTCTCTTTTGACTTTATTTATC
CCTCGAACGGACATTCGAGTAAAAACCCAAAATAGCCTTCATAAATCCAAAA
ACAGGCTCCACATCAATTTTTCTTGACTGTAGATGCTTTTTGTTTTTGGTT
CTGAAAACTTTTTAATAATTTGGACTTTAAAATATTCCCAATTATAATTTT
TCATTATTTTTTTATTTGCTTTTGAATCGATTAATTGTTATATAAGAAGGC
TTTTGATTTTATGATAGACACATCATTCGAATATTATCATTAAGTATTTTT
TCTATCTACGTCCTGAAAACACAGATTGGGTATAGGCATATAAAACTACTT
TTAACATTATTTTTGGATGGTATGATATTGCATCACGATGATGTCTGAATT
CAGTCTCTAGAATTGTTTCTACAATATCATTTACATGTCGTGAAATATCAT
TTGTGGGGATAAGAACTGAAGTTTCCATTGGTAGAGTAAGTTGAGTCATGT
TATAATCTTTATACATAAGGCACCTCGTTAATTTAGTTTAGTGATGTTTAT
TAAATTATACGAAAGGGCCTTATTTTTTTTAAAGTATTTTAATGTAAAATTA
CATATGAATACAAAGTATTTTGGCGAGACTCTTGAGGGAACAGGACAAGCT
GAAGACTACAGGCTGAAGCTGTCCTCTAAGAAAGCGAGCCAACAATACGAA
GTATTGTAAATAAAGAAGCCAGTAAATGAATTTATGAAAACTCATTTACTG
GCTATTTTGCTAGGAATTATGTCTCAGGCTCTTTCCTACGATTAGTCAAGT
ACTTATAATTAAAATAAGCATTACTTAACTAAACCGTTCATTTTAATTTTT
TTGAATAAGGTGTTAAACTATATTTGATTGTACGTTTGGCTAATGTGGTGA
CATTTGATAATCGTACAATTTATGATTCATTACAAGATAATGAATTGTTTT
TAATAATCGATTTATACAAGCAATGATGGCAGTCTTATGAGGTTTCTCATT
AGGCTGCTTTCTTAGTTTGTAGTAATAATCGACGACATGATTGTCATAATG
ATGCTGCCCTCTTATTATATTCATAATCACCCAAAATAAAAGTTTTCTCGC
TTTTTTATTACCACGCTTGTTGATGGTATCTCTACAGTGTGTATGACCTGA
TTGATATCGTTTGATATCAATGCCAACAAAAGCATTGAGTTGTTTATTTGA
TTTAAATCGCTTAATATCACCAATCTCCCCAATAATCATAGCTGTGCTTAG
CTTACCAATACCAGGTATCGAATGAATATTTTCAAAATAATCGAGTTGTTG
TGCTAATTGAATCATGGCATCATCTAATTGTTTGAGATGATGAATAGATTG
TTTTAATTGTTGAATAAGTAAGCGTAATTTTTCGACTAGAAAGGAATGTCT
ATCGACATTAGGATAGCTTTCTTGAGCAATCACCCTTAATTGAAGTGCATA
TTTTGTAGCTTTATCCATTGACATTCCCTTATCTGTAGAATTGAATATATG
TGTAATCAGTACCTCCTTGTCGATATCAAGAACCATGTCTGAATGAGTAAA
GATTTCTGCGATGTTGAGTGCAATGATTGAATATCGACTACTAAACAATCT
TTCTAAACCAGGGAATGTTTGATGGAGTAATTCAAGGATCTGAAATTTAAG
TCGATTTTGTTCATTCTCGATTTCTAGATGAAAACGGACGCGTTCTCTTAA
TTCAAAGAATATTAACTCATGTATAGGTAAGCTGTCTGTTTGTTTAAGCGT
CGGTCCTAAACAAGCAAGCTTATGAGCATCTGCCTGATCAGTTTTCCATGA
TCTTAGAGCGCTCGTTTTAAATTTGGCTTCTAACGGATTCATTTGAATATA
GTTAATTTGATTTACACAACAAAATCGTTCCATACCTCTTGAATAGATACC
TGTAGATTCAAAAATGAGTTGTGGGTGGTCTAAGTCATTCAAATACTTGAG
TAAATAATTGTAACCATTTTTATTATTCTGGATGAAAAACTCTTTTTGGAA
TTTTCCATTTTTATAATGTGCAACTACACTACTTCTTTTACTAATATCAAC
ACCTAAGTAATCGATAAAAAAACCTCCTTTGAATAATTGAGAAGCTAAAAA
CTTTACTTAACCTTTCTCATTTCATTTTCCTATACACGGTTTCAAGAACCC
AACATACTACAAACGAATTTCAAAAGGCGAGAGTAAAGCTGACTTGTTTTT
TATACGGATTTAAAATCCAAGAGTCTGGACAGTCTACTTCTCTCTATAACT
ATAAAAAATAGCTATGAAAAAATCTATCGTCATAGATTTCTTCATAGCTAA
TCTTAGTATGTTTGTAATATATTTAAATTATTGTTCGTCTTGCATTAATTG
TTTAACTCTTTTTGCTTCTTTTTCACGGGCAGATTTATTAAGTATTTTCTT
TCTTAAGCGAATACTTTCAGGAGTTACTTCCACCAATTCATCATCATTGAT
AAATTGTAACGCTTCTTCTA
```

```
Sequence 3367
step.1002b02.cons.ok
CTAAAAAATAAATTGTATATAACTTACTTTGAAATTGATCGAATATATAGT
```

```
ATCTTTAATAAAATGCAGATATTGTGGCGTAGAATTTGAGAATCAAAAAAT
GATTAATATAGTTTGAAGAAACTGAGCATAAGTACTAGAAAAATAGCCAGT
AAATGAGTTTATTATGAATTCATTTACTGGCTTCTTTAAATAATGATTTAG
ACAGTTTGCTTTTCTAGAACTACTTTGAAATGCGTTATTTTCTTTAAGAAT
ATTAATATATAACTTGTGTTGCTAATGCTAACCTTGGTCAGAGTGTATTGT
CATGCGATAGTTAAGATGAGATCTTTTAATTAAAACCTTGTTTTAAACATC
GATGACAAGGTCTAATGTAGGAGGTGTAGAAATAATTTTCAGAATGATATA
AATAAATCCATAAATGAAGATAAATCGAAACTATTTATACCGTAGGGGGGT
ATATGTTATGGTATAAGTAACTTTACTATCACGTTTTCATTGGGAGGGGTT
AATTTGAATAATAATGGTGAAGAGCATAATCATCAAAATCACATGAATCAT
TCCAATCACATGCATCATGATAACCATGCCTCACATCATCATAGTGGCCAT
GCACATCATCATGGAAATTTTAAAGTTAAGTTTTTTGTTTCATTAATTTTT
GCAATACCTATCATTCTCTTATCGCCAATGATGGGTGTTAACTTACCTTTT
CAATTCACATTTCCAGGTTCTGAATGGGTAGTGTTAATATTAAGTACAATT
TTATTCTTTTATGGTGGTAAACCGTTCTTGTCTGGTGGTAAAGATGAAATT
GCTACAAAAAAACCAGGCATGATGACCTTAGTTGCTCTAGGTATTTCAGTA
GCTTATATTTATAGCTTGTATGCTTTTTATATGAATAACTTTAGTAGTGCA
ACTGGTCATACAATGGACTTTTTTTGGGAATTAGCAACCTTAATTTTAATT
ATGCTATTAGGACATTGGATAGAAATGAATGCTGTCGGAAATGCTGGAGAT
GCTTTAAAGAAAATGGCAGAACTATTACCTAATAGTGCTATTAAAGTTATG
GATAATGGCCAACGCGAAGAAGTTAAAATATCAGACATCATGACTGATGAT
ATCGTCGAAGTAAAAGCCGGAGAAAGCATTCCAACAGATGGTATTATCGTT
CAAGGACAAACATCTATAGATGAATCCCTAGTCACTGGAGAATCTAAAAAA
GTACAAAAAAATCAAAATGACAACGTCATCGGGGGTTCTATTAATGGGTCT
GGAACAATACAAGTCAAGGTTACAGCTGTTGGAGAAGATGGATATCTTTCT
CAAGTTATGGGACTTGTTAATCAAGCACAAAATGATAAATCTAGTGCTGAA
TTGTTATCTGATAAAGTAGCGGGTTATTTATTCTACTTTGCTGTAAGTGTT
GGCGTGATTTCTTTTATTGTCTGGATGCTCATTCAAAATGATGTTGATTTT
GCATTAGAACGTCTTGTAACTGTGTTAGTCATTGCTTGTCCACATGCTTTA
GGCTTGGCAATACCTTTAGTCACTGCACGTTCTACTTCAATTGGTGCACAT
AATGGTTTAATTATTAAAAATAGAGAGTCTGTAGAAATAGCTCAACATATC
GATTATGTAATGATGGATAAAACTGGTACTTTAACTGAGGGTAACTTTTCT
GTGAATCATTATGAGAGCTTTAAAAATGATTTGAGTAATGATACAATATTA
AGCCTTTTCGCCTCATTAGAAAGTCAATCTAATCACCCATTAGCTATAAGT
ATTGTTGATTTTGCGAAAAGTAAAAATGTTTCATTTACTAATCCACAAGAC
GTTAATAATATTCCAGGTGTCGGATTAGAAGGTCTAATTGATAATAAAACA
TATAAAATAACAAATGTCTCTTATCTTGATAAACATAAACTTAATTATGAC
GATGACTTGTTTACTAAATTAGCTCAACAAGGTAATTCAATCAGTTATTTA
ATTGAGGATCAACAAGTCATTGGCATGATTGCTCAAGGAGATCAAATTAAA
GAAAGCTCAAAACAAATGATAGCTGATTTACTATCAAGAAATATTACACCA
GTCATGCTTACAGGTGACAATAATGAAGTGGCACACGCTGTCGCAAAAGAA
TTAGGTATTAGTGATGTTCACGCACAACTCATGCCAGAAGATAAGGAAAGC
ATTATAAAAGATTATCAAAGTGACGGTAATAAAGTCATGATGGTCGGAGAC
GGTATCAACGATGCGCCGAGTCTTATAAGAGCCGATATTGGTATAGCAATT
GGTGCAGGCACAGATGTTGCAGTGGATTCAGGTGATATCATACTTGTTAAA
AGTAATCCATCAGATATCATTCATTTCTTGACTCTTTCAAATAATACTATG
AGAAAAATGGTGCAAAACTTATGGTGGGGTGCAGGTTATAATATTGTTGCT
GTACCTTTAGCAGCTGGCGCATTAGCTTTTATCGGGTTAATATTATCACCA
GCTGTAGGAGCAATATTAATGTCTTTAAGTACAGTTATAGTAGCGATTAAT
GCTTTTACATTAAAATTAAAATAAAAGAGGTAAACCTTATGTATAATAAAG
TTTTTGCAATTTTAATTATAATTTTTTCCATAATAATTATTGCGTCTAATG
ATACTTTCGCAGAAAGTAAGAATGATATGATGAATATGAAAGAAGATAAGA
AAAATACAATGGATATGACAAATATGAAACATCATGACGAAAGAAAGAAAT
TAAATTCTTCACAAGGAAAAAATGAAATAATATTTCCTAAAGTTGCAGAGT
CAAAAAAAGATAACAATGGTTATAAAAATTATACATTAAAAGCTCAGGAAG
GAAAGACAGAGTTTTACAAAAATAATTTTTCTAATACTCTAGGCTACAATG
GAAATTTACTTGGACCAACTTTAAAAATTAAAAAAAAGGAGATAAAGTTAAAA
TTAAGTTAATAAATAACTTAGATGAAAATACAACATTTCATTGGCATGGAT
TAGAAGTAAATGGAAAAGTGGATGGAGGGCCTTCTCAGGTTATAAAACCAG
```

GAAAAGAAAAAACTATAAAATTTGAGGTTAATCAAGATTCTGCTACGTTAT
GGTATCACCCCCACCCCTCTCCAAATACAGCTAAACAAGTTTATAATGGCT
TATCAGGATTATTATATATAGAAGATAGTAAAAAGAATAATTATCCTAGTG
ATTATGGAAAAAATGATTTGCCTATAATAATCCAAGATAAAACATTTGTAT
5   CTAAAAAATTAAATTATTCAAAAACGAAAGACGAAGATGGCACTCAAGGTG
ATACTGTTCTTGTGAACGGAATAGTAAACCCCAAACTGACAACAAAAGAAG
AGAAAATACGTTTGAGACTTTTAAATGGTTCTAATGCTCGAGATTTAAATC
TTAAGCTAAGTAATAATCAAAGTTTTGAGTATATTGCTTCAGATGGCGGTC
AATTAAAAAAACGCTAAAAAAATTAAAAGAAATTAATTTAGCTCCTTCAGAAA
10   GAAAAGAAATAGTAATAGATTTATCTAAAATGAAAGGCGAGAAAATCAGTC
TGGTTGATAATGATAAAACTGTAATTTTACCGATTAGTAACAAAGAGAAAA
GTTCTAACAAAGGTAATACACCAAAAGTAAGTAAAAAAATAAAAATTAGAAG
GTATGAATGATAATGTTACCATTAATGGTAATAAATTCGATCCTAAAAGAA
TAGATTTTACACAAAAGTTAAACCAGAAAGAAGTATGGGAAATTGAAAACG
15   TCAAAGATAAAATGGGTGGTATGAAACATCCTTTCCACATCCAT

Sequence 3368
step.1002b04.cons.ok
20   TGATAAATTAGTGGTTAGCTATATTTTTTACTTTGCAACAGAACCGTTTTT
TGTATATAGTCTTATAGGAATTGCTATATTTTTTGTGCCGGTAACGATTAA
TAATAACTCTTCAATTTTATTAGATCATTTTGTCACATGGATAAGTAGCGC
ATTACCTCTTTTAACTAAGATATTCATAATGATTATCATTATACTAGGTGC
TATTTATCCATTTATTAAAGGGACATGGAATCGGAATACCGTTGAAACAAT
25   TTTTAGTTTATTTAAAGTTTTGGGAGTCATTATAGGCGTTTTGTTAATTTT
TAACATTGGGCCAAGTTGGTTACTTAATGAACAAACGGGAATGTATGTTTT
TAACTATTTGGTAATTCCGGTAGGATTAACAGTACCTGCAGGAGGCGCGGT
ATTAGCTTTATTAGTAGGATATGGCTTATTAGAATTTGTAGGTGTTTATGC
GCAAAAAATTATGTACCCGATATGGAAAACGCCTGGACGTTCAGCAGTTAA
30   TGCTTTAGCATCTTTTGTTGCTAGTTTTGCTGTGGGTTTACTTATAACGAA
TAAAGAGTATAAAGAAGGTAAATTCACGGAAAAACAAGCTGTTATCATAGC
AACCGGCTTTTCTACAGTTACTGTAGCTTTTATGATAGTTATTGCTAAAAC
CTTACACTTAATGGATATATGGAATTTATATTTTTGGTCTACCTTGTTTGT
TACTGCTGCAGTAACAGCTTGTACAGTTAGGATTTGGCCTATCAGTAAAAT
35   TAGCAACACATATTATGATCAGCCATTTATAGAAGAAGATACAAGCGAATT
AAAAGGTTAAAAAAAGCTACTTTTTGCATGGGAAAAAGCAATGGAAACTGT
TGAATCATCACCTAACGTGATGAAAAATATTTGGTTGAATTTCAAAGAAAG
TCTGATTATGACTATGAATATCTTACCCACCATATTATCAATAGGTTTAAT
TTGCTTGTTACTCGCAGAATATACAGTGATTTTCGATTATTTAGCATATGT
40   TTTTTATCCATTAACTTGGATACTTCAAATACCAGATTCCTTTTTAACTGC
AAAAGGCGCAGCTATTGGTATAACAGAAATGTTTTTGCCTTCATTAATTGT
AGTCGAAGCACCATTAATCACTAAATTTATAATTGCTGTTACTTCTGTTTC
TACAATTATATTCTTTTCAGCTAGTGTGCCTAGTATTCTCTCTACTGATAT
ACCCATCCGCATAAGAGATTTAGTGGTTATATGGTTTGAGAGAACTGTATT
45   GAGTTTAATTATAGTAACACCTATCGCATATATTTTTTTATAATTCAATAA
TAAAATATCGTATATAAAAAAAATCACCTTTTTTTACAAATATTTATAGTGAT
AAGGAAAGCAAGGATATGAAATTTCCATCTTAATTTACTGTTTTTTATTTT
AAACTCAATTATATATTTAGGGTACACCCTAATCATACAAAATGAATTGTT
GAAATTACAGCAATTAAAATGTATGATTAGGGTGTATATTAATTATGAGGT
50   GAAACTAAATGATTGTGGGCTATGCGCGGGTATCATCAGTAGATCAAAATT
TAGAAAGACAACTTGAAAATTTAAAAAACATTTGGTGCTGAAAAAATTTTTA
CAGAAAAACAATCAGGTCAATCTATAGAAAACAGACCCGTCTTTCAAGAAG
CATTGAACTTTGTGAGAATGGGAGATCGTTTTGTCGTGGAATCAATTGATC
GTTTAGGTCGTAATTATGATGAAGTCATTAATACCGTTAATTATCTTAAAG
55   AGAAAGAAGTACAATTGATGATTACCAGTCTACCAATGATGAATGAAGTCA
TTGGTAATCCATTATTAGATAAATTTATGAAAGATTTAATTATACAGATAT
TAGCAATGGTTTCAGAACAAGAAAGAAATGAAAGTAAACGTCGACAAGCTC
AAGGGATTCAAGTTGCGAAAGAGAAAGGCGTATATAAAGGACGACCTTTAT
TATATTCTCCCAATGCGAAAGATCCTCAAAAACGTGTTATCTATCATCGTG

```
         TTGTAGAAATGTTAGAGGAAGGCAAAGCGATTAGTAAAATTGCGAAAGAAG
         TGAATATTACAAGACAAACTGTTTATAGAATTAAACACGATAACGATGATT
         TGGGTTCTGTTGCAAAGTAAAAAAATATAGCTAACCACTAATTTATCATGT
         CAGTGTTCGCTTAACTTGCTAGCATGATGCTAATTTCGTGGCATGGCGAAA
    5    ATCCGTAGATCTGAAGAGACCTGCGGTTCTTTTTATATAGAGCGTAAATAC
         ATTCAATACCTTTTAACTTCGAAATGTTTTTTTATTTCTTCGGCAAAATAT
         TGGGGTGTTAAAATATTCGGAGGCATTTGACTTAGTTCACGAGCATAATTT
         ATAGATTCACCCAATTGTTGGCCTTCGTTAATCTTCGTTTCTATATTATGC
         TCTGTCGTTTGTATATAGACTCTACCTTGATATGGAGGCTTTTTATCCGAT
   10    TTATAATGGTCAAATTCAAAGATAGATTGTGCACTTTGTAAACCTAAAGTT
         TGACAAATTTCAACAATGTTACCATGTTTAGATTTGAAAGTATCTAACAAT
         AATTCACAATCCGTGACATGAACGTCTTTTAAATATTGAAACAACTGACCC
         CAAACTTTTAATAAACGTTGATATGATAAGATTTTTAAGTTTCCTAGTCCT
         ACAGTTAATAATCTTTTAGGCTTTTCATCAATGTAGATAAGCGTTGATGAT
   15    ATTTTGCCGATTGAGCTATTGATGACATGATATTGTTTTAATGTTTCTAAT
         CGTTCTTTAATTAACGTCTTATGATATTTTATCTCACCTAATTGGTTGAGG
         TGATCTGGCACACCCACCACAATCGTTTCAATTGATGTTGGCGTATCTTGA
         TTAATAAAAAATTGCATATAGATCCTCCGTTCCTCATTCTGTATAACAATA
         ATTATAAGTTATGAACACCCCAATATTAAACAAAACCACTTTTTACTCTAA
   20    CTCTAAATAAAAACCACTTTTGTTTAATTAAGTGTGAAATTCTTCTTCATC
         ATATTTTTCTGATAAAAAATAACCTGGGACAATCATTTTGTCCCAGGTATT
         ATTAAAATATAAACAATCGATGACAAAATTAATTTGAGTTGTTTACATATT
         TTAACTCTAGTCATCTTTGCTAAGGGAGACAACAAATTATGTCTCGTACCT
         TTTTATTGTTTTTTTATATCTATTACTTGATTAAAATTTTCCTTTTTTAAA
   25    GGCTAATCCTATGCCACCAAGTTTGTAAATCGCACGAGTATCAATAACTTT
         CTTCAAGAATGCTGCTTTTTTACCAGCGAT


         Sequence 3369
   30    step.1002b05.cons.ok
         CAATATGTTCGAGTAATGACTTGCGCGCATCGTCTTGCTGTTTCAAATATT
         CAAATAAAAAGCCAGTCCATACCCAAATTGTTTTTGTATTAAGAGCACTAC
         AAGGCGTATTTGCTGCATCATTTTCTCCTGTAGTCATGACTTATACAACTG
         AAACTTATCCACGTGTGAAACGTGTAACAGCAATTAGTTTCATCAGTACAA
   35    GTTTTATGTTATCAGGTGTGTTAGGGCAAAATATGAGTGAATTAGTCGTAA
         GTTATTTGAATTGGCAGTGGGTATATTTTATTTTAACAATCTTATATTTAA
         TTCTCGTATTAGTTATTTATAAAAATGTACCTGAGAGTCCACATAAAAATC
         CTGATATTCAACTCATTAAGTTTTTCAATAACTTTAAAGATTTCAAAGACA
         ATCTTAAAGTTTTCTATTGTTTATTTATTTCACTAACATTACTGATTATGT
   40    TTATAAGTATGTATGATATTTTAAATGAATATGTCACATCACACCAAGTTG
         GGGGAGACATGTCTGTGTCCTCAATGATGAAATTGTTCGGTGTGATAGGCA
         TAGAAACTAAAGTAATCATTTTTATTTAAAAGTAGATTTTAATTTTTTATAT
         ACATATTAGCGTTGTGGAATAGGTGGAAAGGGGGAGTCAAAAGGTGAAAAC
         GCTAATGATTAAAGCAATGGGGACGGTGATACGTTTATCGATTGAGCATCA
   45    ACATCCGGATACATTACTTCAAGAAGCTGAAATAAAAATTCGTGCTTGGGA
         ATCACAATTTAGTGCTAATGATCCGAAATCAGATTTGATGAATGTGAATCA
         GCATGCAGGTATCGCACCAGTCAAGGTTAGTTCTGAGATGTTTAACATGAT
         ACGTTTTGGTTACGAAACTACATTATCTTCTAATTTTAAGATGAACATTTT
         GATAGGGCCACTAGTCAAATTATGGAAAATTGGTTTTAAAGATGCATTGAA
   50    ACCTAAAGAAGAGGATATACAACGTGCTTTATTGTGTATGGATCCTGAAAA
         TCTTGTTCTAAATTCAAAAACACATGAAGTATTTCTTACACAATCAGGAAT
         GGAGATTGATTTAGGAGCTATAGTTAAAGGCTATTTTGCTGATCAATTACA
         GCAATACTTTTTAGCTCATGGTGTATCTTCTGGCATTATCGATTTAGGTGG
         TAATGTTTTAACAATTGGTAGACAACCCGAAACATTAGAAAAATGGCATGT
   55    AGGTGTACGTAATCCATTTCATAAGGATACACTACCACTCGTTACATTAAG
         CGTAGAGCATCAATCAGTTGTCACATCAGGTATCTACGAACGCTACTTCAT
         ACAGGAAAATCAATTATTTCATCATATATTGGATTCAACAACAGGTTATCC
         TGTAGATAATGATATCGCTAGCGTGACAATCATATCTGATCATGGGATTGA
         TGGCGAGGTATGGAGTACAATTTGTAGTTTTGGTCAGTCACAAAAAAAATAT
```

TGAATTATTAAATCTCATTGACGGTATTGAAGGCATTATTGTGACAAGAGA
TGGAAGCGTTTTAATGACTTCGAAAATGCAAAAGTATTTATAAAAATTTAA
AGAATACTCATGAAGAGTATTTTATTTTTAAGTAATAATGTGAAAATATTC
ACAAAAAAACTAGGAGGATTTGCTATGAGTAAGGAAATATTCGATACTTTT
5   AAATTTAAATGTGGTGCCGAATTAAAAAATAGAGTATTAATGGCACCCATG
ACTATCCAAGCTGGGTATTTTGATGGAAGTGTTACATCAGAAATGATTGAT
TATTATCAATTTAGAGCTGGTGATGCTTCAGCAATCATTGTTGAAAGTTGT
TTTGTTGAAAATCACGGACGAGGATTTCCGGGAGCTATAGGTATTGATAAT
GATGACAAAATACCTGGACTCAAACGTTTAGCAGAAGCGATTCAAGCTAAG
10   GGATCAAAAGCGATTTTGCAACTTTATCATGCCGGAAGAATGGCAAATCCT
AAATTTAATGAAGGAGAGCAGCCGATATCTGCGAGCCCCATTGCAGCATTA
AGACCTGATGCTGTACCACCTAGAGAAATGACACATGCTCAAATCAATCAG
ATGATTGATGACTTTGGAGAGGCTACACGTCGCGCTATAGAAGCGGGGTTT
GATGGTGTCGAAATTCATGGCGCCAACACATACTTATTACAACAATTTTTC
15   TCTCCACATTCTAATCGGAGACAAGATTCATGGGGAGGCAGTCGTGAAAAA
CGTACACGATTTCCAATCGAAGTTTTGACAAAGGTTCAACACGTCGTTGCT
GAAAAAGAGGCTTCTCATTTTATTATAGGATATCGATTCTCACCTGAAGAA
ATTGAAGAACCAGGCATACGTTTTGAAGATACCATGTTTTTACTAAATACA
TTAGCAGAATATGAACCTGATTACTTCCATATATCAGCAAACAGTTATCAA
20   CGTACATCTATTGTGAATCAAGAAGATACAGAACCTTTAATTAATAAGTAC
ATCAAAATGCAAAGTGCACAGTTGGCAAAAATTCCATTAATTGGTGTAGGT
AGTATTGCCCAACGACAAGATGCAGAACATGCCCTTGAACTAGGATATGAT
CTTTTAAGTGTTGGGAAAGCCTATTTAGTGGAACCACAATGGACAGATAAA
ATTTCACAAACGAAGAAGTAGAACAATTTGTCGATATACATGATCAGAAA
25   GTACTTCACATACCATCCCCTTTATGGAAAGTAATGGACTTTATGATTTTA
GATAAGAAGAAGAGCATCGTAAATATGAAAAATTAAAAGCACTTCAAAAT
AAAAAAGTTAAATTTAACAAAGGTACGTATCATGTCTATGCAAAAGGTCAT
AATGGCAACTTACCTATGAAAGTCCAATTATCAGAAGATAAGATTGTAAGT
ATCGAGGTAGATGATAGCGGAGAGTCTGAAGGCATAGCGAACCCAGTGTTT
30   GAACGTTTACCTCAAGATATTATCAATGGGCAAACACTGAATGTAGATGTC
ATTTCAGGTGCGACAGTAACAAGTGAAGGCATCGTGCAAGGTATTGCAGAT
GCAATTGAACAAGCAGGAGAAGACCCAGATATTTTACGGGCGCGTCCTAAA
CCAGTCGTTCAGTGGTCTGATGAGGTTGTTGAAGAGA

35

Sequence 3370
step.1002b07.cons.ok
GATAAGAATTCTTAATTCAACTTCGGCAATATTATAGTCTATAATATACAA
GTCAATTTATTCAGTTGTTTCAATCAGACCGTATTTGCCATCTTTTCGTCT
40   ATAAACAATGCTTGTGCCATCAGTTTCTCTATCGGTAAAAATAAAGAAATC
ATGTCCCAATAAATTCATTTGCAATACTGCTTCTTCAGAATCCATTGGTTT
TAAGCTGAATTGTTTAGAACGAATAATTTCAATATCATTTTCACTTTCTAT
ATCATCTGCATGATTGTTTGTTGTAGACTCTTGTACTTCAGCAACAAAGAT
ATCTTGATCTCCACGATCTCTATGTTTACGATTTACACGAGTTTTGTATTT
45   ACGTACTTGTCTTTCCAATTTGTTAGTAATTAAATCAATGCCTGCATATAA
ATCATCATGTCTTTCTTCAGCTCTAAGAGTGACATCTTTTAAAGGAATTGT
AACTTCAATTTTAGTTGTAGAATTAGAATAAGTTTTAACTCTAACATGTGC
TACAGCATTTGGCACATTGTTAAAGTATCTTTCTAATTTACCTACTTTCTC
CTCAATATAGTTGCGAATTGCATCTGTGATAGTGAGGTTATCTCCATGAAT
50   TTCAAATCTAATCATAGTAATCTCTCCTTAAACCTCTATATTGATACTTCT
TACCTATATTTTACCATGTTTATGTTATCGTGCAAACGCAAACACTTTGAA
TTTTCTGACTTTTTTTATCGTACAATTTACACCCTGCGTGATGAATAGTTAA
CCCTGTTGTATAAATATCATCTATGAGTAGTATTACTTTCCCTTCGATATT
TATCTCTTCATCTTTTATATAAAATGGATTAGGGGCTTTTGAACGTTCAAT
55   CTTTCCTAACTTTGACTGCTTAGGACGTATATGTGTACCTAATACATCTTG
ATATGAGATGCCCATTTTATCTAAGACAGTGGTCACAGGATTAAATGTACG
TTCAATGTCGCGTTCAATTGGAGAAGGAATGGGAACTATATAATCATATTG
CGTTTGAGGTAAAACTAATTTTCTCGCCAATACTTCTGCCAAATAGAAGTC
TCGCTTAATCTTATACTGGTGTATGATTTCTTTCATCATACCATCATAACT

ATAATCACAAAACAATTGTTCCATTAAATAAAATGTTTTACTTAAGAATTG
ACAGTCATGACATTCGCCATCAGTAGAAATTTTGTTGTTAAGGCACTTAGG
ACACCTACCATCTTTATGAATTTTGGAGTTAAGCCATAATGTTTCACATCG
TTCACAAAATATCTGAGGTGGTTTAAATAAGTTATAAATATGAATGGTCTC
ATGAATAGGTTGATTGCACTGGATACATTTACGCATCAATCCATCCCCTTT
TAATTGCTAAACGATTCATTGAAATAATGTTTCTTTTAGCTAAAATCATCG
ATAATGTAACACCTTCATGAAGAAATAAAACTAAGCCACTTGGAGACTGCT
GTTTACGTCCTACGCGACCAGCAATTTGAATTAAAGCCTCTTGTTGAAAAC
TTCCAGCATCAACTACAACGACATCTAAGTGTGTCATTGTAAATCCTCTTT
CTAAAATAGTTGTAGTGAATACAATTTTGTGTTGTCCTCGTCTTAAAGCTT
CAATTTTTTCAAATCGTAAATCATCTTCACTGTGAACGCAAATCAAATCAG
GGATGTCCATTTTATACTGTTGATACATTTTATTCATAATTTCTATATTAT
TAATAAAGACCAAAGTAAAACGTTGTTGGTTAATTTGATATCTAAATATAT
TAAGTAATAAATTTTGTTTTCGTGTTGATTTTAATTTGAAATATTTGAACT
TAGGAATAGGAAGGGGGGATCGGTGAAAACGGGCTGGTAACTTAATTATTT
TTTCTGGGGGGAATTGTTTTAAAAAATGACGCGGTGGTGTGGCCGTCATGA
AAATATGTGAATGATTCGATTTTGAAGCAAGTTGTATTGCATTTGATAATT
GTGGATCCATAGACAACGGAAAAGCATCTACCTCATCGACAAATACAGTAT
CAAAATGCTGTTTAAACCTCAATAATTGATGGATAGTAGCAATAACAAAAT
GACCATTATATTGCTGTCTACTAGATTGATGTAGCACATCTATATGTTCAT
CGATAAAAGCATCTTTAATTCGATGACTTATCTCAATAATAACGTCTACAC
GAGGTGATACAATAGCAATATTATGTCCCATCTGACGAGCTATCCTAATTC
CCTCAAACATCATTTCAGTTTTCCCTGCTCCTGTGACGGCATATAACAAGA
GGTCATTATTATTTTTTATAGCTTGAACGATAGTTTCTGAAGCGTATTGCT
GTTGCTTTGATAATTCAAAGGGTAACTGATAATTCGCTTTTGTTGCTTTTT
GAAAACTTTTTACGAGACATATATCTGTAATACTATCCATTCTGCCTAATT
GAATGCATCGTCTACAGTATGTGATTGTTTTATCCAATATCGTTGAATTGT
AAGTCACAAAATATTGCAAGGAAGTGTTTCCACATTGAATACAATACCAAC
GCTCACTGTCTTTTATAACACCTTTTGAAACTTGAGATACTGATTCATCTT
CAAGTTGATGCGTTGGAGAAATCAGTCTACCGTAATATTTAATTTTACTTC
ACTACTTTCAAAAAAGTCTGAAATACTCAAATAAGCGAACACAAGTGATAT
CTAATGATATTACCACTGTGTTCTTTCCTATCAAGCTTTCAGACTTTTGAT
TATATATTAATTAATGTCGATTCTTCTTGGGAAGTAACCCAATCCTAAACC
GCCTGATCCTAAATGTGATGCTATCACTGGTCCAAATTCACAATACTGAAT
ATGAACATTAGGATGATCTTCCTTTAATTGCTGAAGAAATGACTTTCCATC
TTCAGTTTTATCACCGTTTATTACAAATACTGTCACATCTTCCATGCCTTC
TATTTCTTTAAAAATGTTTGTTTCTAAAGATTTTAGCGCACGTTTTTTAGT
ACGTACTTTTTCGTGTGGATGTATTTTACCATCTTCTTCAAAACGCAAGAC
AGGTTTCATTTTCAATAATGTACCTACCCAAGCTTGAGCTCCAGTGATACG
ACCACTTTTTTGTAAATTTTTTAAATCATCAACAATTAAGTATGCACCAAT
ATGTTGTCTTATTTCAGTTAGTTCATTAATAATATCATCAGGTTTATATCC
CTTTTGTACCAATTGAGCAGCGTAAATTGCAAAGCTACCTTCAATCATCGC
AGCAAGACGGCTATCAAACGTATGTACTTGAATATCTTCAACCATTTCACC
AGCTTGTGTTGCTGAAGGATAGCTTCCGCTTAT

Sequence 3371
step.1002b08.cons.ok
ATGTACATGTTGGTGTGTTTGACGTTCGATATTTTGTAATCTTTGGTAAAC
AAGCTTCCATGTTCCACATGCTTCTTCTGGGGTTTGATAATCAGAGGTTAG
TTGATTCATAACATTTAATAATTGCCCAGTATCTTCGTGGTCATTAATTAA
ATCTTCTAATATGATTTTAATGTTTTGTACATCTTGTCCTTGACTATATTG
AATGAGTTTAGGAAAATCTTCCTCATCTTCTTTACGTATATGATCCAACAT
ACTATCACGAAACTCGCGATATAAGTCTTGTAATTTTAATAAGTATGGGTG
ACTAGGACCATGTACTTTTGCCAATTTAGTCATGTAAGGTGTAAGATTCTT
AAATTCTTCTTTAAGCGTTTCGTGATAAGCTGATTGTATATATTGTATAAG
AGATTCAACATTTAAAAAATTTAGGATTAATGGTACTGTTACCTTCTGTATT
ATCAATATGATTCAACTTATTTAATAAGGAATTTAAGTCAATATTTGGTTT
ATGATTGACAGCTGAAGCGATACTCTCTTGTCCGCCACAACAAAATTCCAT

```
         TACGAAGATGTTGAAAAACGTAGGGAAGAATTTCATCCTTCAGGAAGTCGT
         TCTAATTTAACCATTAAAGCAATGGGAAAAGAAATTCTTCCTTATATCGAT
         GCGACATTTCCAACTTATAAAGTAGGTAATACAAGGTTACTTATTGGAGAT
         AGTTTAGCAGGAAGTATCGCTTTAATGACTGCAATGACTTACCCAACTATT
   5     TTTAGTCGAGTTGCGTTATTGAGCCCAATGTATAATGAAAATATTAAGAAA
         AAAATTGATACATGTATGAATAAAGGTCAATTGACGATATGGCATGCCATT
         GGTTTAGAAGAAGCAGATTTTATTTTACCAACTAATGGTAAAAGAGCTAAC
         TTTTTAACACCTAACCGTGAATTAAATCAACTGATTAAAGAAGATAATATT
         GAATATTTCTATAAAGAATTTAACGGTGGACATCATTGGAAATCATGGAAA
  10     CCATTGCTAGGAGATATTCTCTTACAATTTTTAGGTGATCCAATAAATGGA
         AAATATGTTTAATAAAGTAAGATAAATAAAATAGTATGTTAGCGAATGGTT
         GTTGAATGCGAAACACTATTAATTTCAGAAATGTAATTGTTTTCATGATAA
         AAGTAACGGTTTTCAAAAAGTTTTGATATAATAGGAATAAGTTAAACAAAG
         GAGGAATTTAAATGATTTTAGGATTAGCATTGGTTCCGTCAAAGTCATTTC
  15     AAGATGAGGTGAATGCTTATCGCAAGCGATATGACAATCATTATGCTCAAA
         TAATGCCTCATATCACGATTAAACCTCAATTTGAAATCGATGATCATGATT
         TTAATTTAATTAAAAATGAAGTGAAAAATCGAATTTCTAGTATTAAACCAG
         TAGAAGTACATGCTACAAAGGCATCTAATTTCGCTCCAATCAGTAATGTTA
         TATACTTCAAAGTTGCTAAAACAGAGTCATTAGATCAATTATTTAATCAAT
  20     TTAATACAGAAGATTTTTACGGTACAGCTGAACATCCTTTTGTACCACATT
         TTACAATTGCCCAAGGTCTAACAAGTCAAGAATTTGAAGATATATATGGTC
         AAGTAAAATTAGCAGGGGTAGACCATAGAGAAATAATTGAAGAACTATCGT
         TACTTCAATATAGTGAAGAAGAGGACAAATGGACTATTATTGAAACTTTTA
         CATTAGGATAAAAAGTTCAAAATTGTAAAGTGAAATTGAATTTACAGAATC
  25     ATTATTGTTAAATACGTGAAGAGCGCTTAATCAACTAAAAACGCCAAATCC
         TATTGTGTTTCAGTGGGATTTGGCGTTTTTATATGAAACATAATTATATTG
         TATAGTTAATCATTCAAAAAGATTAATGTACGTTTTATATTGAATATTCAT
         ATTTCAAATGCATATGAATGACGTTTCATTAATATATACAACTAATCTTTA
         TATAAACGATTTATAGATTTTAAGGTTTAACCCTTCTGTTTGTTTTAATAA
  30     AGTAATATCCGTAAAATACTGCAAGTGCAAGAAAAATCATCGCTGAAAAGT
         AAAAGGTATTATTTAAATTATTAGTAAATTGAGTAATTAGACCTCCGACTA
         GTGGGCCTATCATTGAACCGAAGCCTTGAACACTGTTGAACACGCCCCATG
         TTTCTTCCTGTTCGTTAGGATTAATATGCCCAGCCATAAAGGTATTCCAAG
         CCGGTAAGAGGATACCGTACATTAGCCCAATAAAAAGTCCTATGGCCCAAA
  35     CTATATATATATTTGTAATTGTAGATAGCCCGAATATAAGAATTGTATATA
         GTATAAAGCCACTAAAAATAACTCCATACATAAACCCTTTGCTATTATTGT
         CGATGATTTTTGATAAAAATAACATAGAGAAAGCACAGCCTATGCCACCAA
         TAATGATTGCTACTGTATATTCAACGGTTGATACTTTCACAACTTGCGTTG
         CATATTTTGGAAGAATAGGTACAAGTGCTGCTATAGCTGCTCCTTGTAACA
  40     AGATACCCGGAAATAGAATAAGATGACGTTGTGTAACATCTACAATTTGCT
         TTAATTGTGCTTTCACAGGTTTAGTATTGTAATTTGTTAAGTTGATATTAA
         CAAAATAGTATAGTACCCAGGCAATAAGCACAACCAAGGCCATTAAAAATG
         CAAAACGAGTAGGATGGAATTTAATAAGCAAGTTCATGATAACCATACCCA
         CCAATAAACCTAGCAACCATGAAAAGTAGACATAACCCATTTGTTTGCCGC
  45     GATTTCTTTCATCTACACTAGATAACATGATAACCCAAATCGGACTCACAG
         CAATTCCTAACATAATAGCACTGAAAATAATTATGAATGGTGATGCCGGGA
         ACCATATCACTAAAAACAAGCTTATAAAAGCAAGTAAG
```

```
  50     Sequence 3372
         step.1002b09.cons.ok
         TTGCTAATAAATATGGCAAGGTTAAGGCGATTGCGAGTGAAATGGTTCCAT
         GTACACCACATAGAGTCATGATTAGTGCATATAAGCTACGCTTAGGTTTAC
         TCTCAGTAACTTTATCTTCATCATTTTTAGAAATCATTTTTTGGAATGGAC
  55     TGACAGATAAATAGAAATAAGGATAAAGTACATATACCCACACAAATCTAA
         AGAGATACACCGCAAGAGCGACTAGCAATGTGATGACGATTAAGAATAATA
         AATTATGCGGTTCAGTTTTAATAATTTTCACTATTACCTCAGGTACTAAAA
         AACCTAAGATTGAAAACACAAAACCATTTAAAACGTAGCCTAAAATACTCC
         ATGTGTGATTATAGCTCATCTGTAATTGTGTTCTCGTTTGAGCTATACGGT
```

```
        CACGTTCAAATCCATGCACGAGTCCCGCAACTACTGCTGCAATAATCCCTG
        ATGCATGGAATAATTCAGCAATCAAATAAGTGACAAACGGCGTCAGCAATT
        GTATAAATGTAAACATGTTAATGTTCTCAATGCCACGTCTCATTAATGTTA
        AACGAAATCTCACTAAAGCCATGCCTATAATTAGACCAACAATAGCACCTC
5       CAATAGAGGCAATCAAAAATTGTTGAATCGCATCAAAAATGGAAAAAGTTC
        CAGTAATCAATACACCTACGGCAATTTTAAATGAAATGATACCTGCAGCAT
        CATTAAGTAAAGATTCACCTTCAAGAATGGTCATTGAACCTTTTGGTAAAA
        CCTTACCTTTCGTGATTGCTTGTACCGCAACAGCGTCTGTAGGACAAAGTA
        TAGCCGCAATTGCGAAAGCAGCACCTATAGGTAATTCAGGCCAAATCCAAT
10      GAATAAACAATCCTACGCCTATAACGGTAGTTATGACAAGACCTAGAGCCA
        TCATCATAACAGGTTTGATATACTTCCTTAAATGTACTCTTGAAACATTAA
        CACCCTCAACGAATAGTAGTGGTGCAATCAATGTAACCATAAATAATTCAG
        AATCAAAATTGAATTCAACTGGAATGGGAGTGAGGTATAAAATCATCCCAA
        GGAAAATTTGGATAAATGCAAGAGGTACTTTAGGTATGAAAGTATGTACAA
15      AGGAGCTCACAATAACTAGTGCGACAAATATGAGTAGAGTTTCAAATATCT
        CCACACTTTCACCTCTTTATTAATTTTTAATTGCAAATTTAGTCATCAAAA
        CAATAAAAAGTTGCAAGGTTATCTACTGCAACACTGAAGAAGTGAAATAAT
        AACAATAGAGTCATTCCTGTAAGGACATAACTATTAACACTCAATCATGCC
        ACAAATATACTATTCAAAACTCAATTATTTTCTTTATCAGTGAAATCAAAC
20      ATGAACATTTAAATTAACTTATGAAGCTCGCTTCTTCGAACGAGAAAATGC
        TTATGTGTAATTTCATTGAAAAAATTAGTGACAACTGTCATTCACAATTAT
        TCTTGTTTACAATTTTATCATTTATTAACCAGATTCTTAAATAAATAGCTT
        TTATATCATTAATATCTTCACCTCAGATATACAAAAAACTACTAGCATATA
        AAAGATATATCTTTTATATGCTAGTAGTTTAAAAGCTTAAACAATCAATCG
25      TCCATTTTATTAATATATTCTTTTTCTTTATAGTATAATTTTTCTCGTCTT
        TCTTCTCGTAGTTTTGCACGTTTTTTTAATTCTTCTGGCGTATTAAACTCT
        TTTTGGCTTAATAAGATAGTAGATTGTCTAGATGCATCGTCTTTCCTGTAC
        ATATATGCACCATTACGGTAGGCCGCTCTACTTATCAAATGCATTCCAACC
        GGAGATGTAAGATTGATAAAAACTAGTGATAATAATAATCTGACACTGAAA
30      AAACCTGAATTCACAATAAAATAGATCAGTACACCAACTACAGTTAGTAAT
        ACTGACAATGTAGAACTTTTCGTTGAGGCGTGACTTCTTAGAAAGACATCT
        TGAAATTTTACAATCCCTATTGCACTAATTAATGCAATAATACTTCCTAAG
        AAAATAAGTATCTTATTATCTAATTAAAAAACAATAGAATACTTAGTTTGA
        CTGTTCATGATTAAAATATTCTTTTAATTCTTCCCACTAAATACGAATAAA
35      AGCATATTTTTACTGTTCTTACGCATTCTTCCGATTAGCTTTTCTGTATCG
        TTTTTCATTCAATAATAATTTTCTTAACATCTTCTAAGTAATCATCTTCAA
        TCATTGGCTTTAAAGTATTTCCATCAATAGCATTTAACATTACATCTTCTT
        TTTCCTTCAGTCTGTTATAAACTGCTTTATCGATAAACCTTTATGGGCTC
        TATCCCCTTCAGACATCAAATAATAATACCTTGTATATTGATTATTATTTA
40      ACCCTAAACGATGTATTCTATCACGTGATTGCAACATAAACGTTAAATTAA
        AGTTATATTCAAAATATATTGCATCATGTACTGTCTGATGTAAAGATATGG
        ACTCGCCTAATGTATTAGGATTAGATATTAGAACTTGTGCATTTCCATTCC
        TAAAATTATTGATCATATCTACCCTATCTTCTTTAGGTGTTTCTCCATAAA
        TCAATATTGAATTAATATCACTTTCTAGTAACCTCTTATTGATTTTATTCA
45      TTGTACCTACAAACAATCCCCATACTAGAACTTTTTTTCCTTGGCTAACTA
        AGTTTTCAATCAATTCAATACCTTTTTCGAATTTTGGAGAAGTTACATTTT
        TTAAATCAAAAGAATTATAAATTTGTTGTTTTACTTGATTTTCTTCTTCTT
        TATTTAATGCTTTATCCAAATCGAAATTTAATTCATCATTCAACATTCCTA
        GTTCGCTATAATTAATATTCTTTTGCAATAATTCTGGGTTAGTTGAAGCTT
50      GTAGTAACCTTATATAAATCGCTAGTATCCCAGATTCATTTTCGTATATCG
        CTTTTGCTAATTCAATTTGAATATTACTAGGCTTAACACATAATATAATAT
        CATTTTCTGCTTGAGGCACTTCCAAATCGTTCTTATTTGTACGCCAGAAAA
        AAGGGTTCAACTTGTCATTAATTTCATTAGGATCAGGATTTTGTAAATCAG
        CAACATTCCAACCAAAATAAGTATCATACTCATCTTTATATAAAATATTTA
55      AGAAGTTAAATATGTCTTGATAACTATTTGGAATTGGGGTACCAGTTAAAA
        CGTATCTGTAATAACTTTTAGGACCTAAAGTTAATGCATAACTTGCTCTGC
        TACTATTAATACCTTTTATCCTATGAACCTCATCATAAACTATCATTGTCT
        TATCATTAATAAGTTCATTTAAAACCCCTACATACTTTTGTATTGCTTCAA
        AATTCAAAACAATGACATTTGCACTTCCCCAATCTGTTCGTACTTTACCTA
```

AATCATTATATTTTTTATCTCTTAGGTTCATAAAGTGTAATTCTCTTTTAT
CTTGAAAAACTTCAATAAATTCTGAACGCCAAGCTTCAAATGCATTAATTG
GAGAAATAACTATTAATTTATCAACTTTCCGCTTTATTTCACTA

5

Sequence 3373
step.1002b11.cons.ok
GCCTTCTTATAAAACGATTAATCGATTCAAAAACAAATAAAAAAAAATGAA
GAATTATAGTTGGGAATATTTCAATGTCCAAATTAATCAAAAGCTTTCAGA
10 ACGAAAAGCAAAAACCATCTACAGTCAAAGAAAAATTGATGTGGAGTCTGT
TTTTGGAATTATGAAGCCTATTTTGAGTTTCACTCGAAAGTCCGTTCGAGG
GATAAATAAAGACAAAAGAGAACTAGGACTTGTGCTAATGACACTTAATAT
AAGAAAAGTACCAGCTCAACGAGCTGAAAATTATAAAAAATCATAAAAAAG
ACAATTTTTATATTATTTCAATAGAAATTGTCTTTATTTACTTATCCTGGA
15 ACTTTATGTTCCGGACTCTTTCCTACGATTAGTCAAGTACTTATAATTAAA
ATAAGCATTACTTAACTAAACCGTTCATTTTAATTTTTTTGAATAAGGTGT
TAAACTATATTTGATTGTACGTTTGGCTAATGTGGTGACATTTGATAATCG
TACAATTTATAATTCATTACAAGATAATGAATTGTTTTTAATAATCGATTT
ATACAAGCAATGATGGCAGTCTTATGAGGTTTCTCATTAGGCTGCTTTCTT
20 AGTTTGTAGTAATAATCGACGACATGATTGTCATAATGATGCTGCCCTCTT
ATTATATTCATAATCACCCAAAATAAAAGTTTTCTCGCTTTTTTATTACCA
CGCTTGTTGATGGTATCTCTACAGTGTGTATGACCTGATTGATATCGTTTG
ATATCAATGCCAACAAAAGCATTGAGTTGTTTATTTGATTTAAATCGCTTA
ATATCACCAATCTTCCCAATAATCATAGCTGTGCTTAGCTTACCAATACCA
25 GGTATCGAATGAATATTTTCAAAATAATCGAGTTGTTGTGCTAATTGAATC
ATGGCATCATCTAATTGTTTGATATGATGAATAGATTGTTTTAATTGTTGA
ATAAGTAAGCGTAATTTTTCGACTAGAAAGGAATGTCTATCGACATTAGGA
TAGCTTTCTTGAGCAATCACCCTTAATTGAAGTGCATATTTTGTAGCTTTA
TCCATTGACATTCCCTTATCTGTAGAATTGAATATATGTGTAATCAGTACC
30 TCCTTGTCGATATCAAGAACCATGTCTGGATGAGTAAAGATTTCTGCGATG
TTGAGTGCAATGATTGAATATCGACTACTAAATAATCTTTCTAAACCAGGG
AATGTTTGATGGAGTAATTCAAGGATCTGAAATTTAAGTCGATTTTGTTCA
TTCTCGATTTCTAGATGAAAACGGACGCGTTCTCTTAATTCAAAGAATATT
AACTCATGTATAGGTAAGTTGTCTGTTTGTTTAAGCGTCGGTCCTAAACAA
35 GCAAGCTTATGAGCATCTGCCTGATCAGTTTTCCATGATCTTAGAGCGCTC
GTTTTAAATTTGGCTTCTAACGGATTCATTTGAATATAGTTAATTTGATTT
ACACAACAAAATCGTTCCATACCTCTTGAATAGATACCTGTAGATTCAAAA
ATGAGTTGTGGGTGGTCTAAGTCATTCAAATACTTGAGTAAATAGTTGTAA
CCATTTTTATTATTTTGGATGAAAAACTCTTTTTGGAATTTTTCATTTTTA
40 TAATGTGCAACTACACTACTTCTTTTACTAATATCAACACCTAAGTAATCG
ATAAAAAAACCTCCTTTGAATAATTGAGAAGCTAAAAACTTTACTTAACCT
TTCTCATTTCATTTTCCTATACACGGTTTCAAGAACCCAACATACTACAAA
CGAATTTCAAAAGGCGAGAGTAAAGCTGACTTGTTTTTTATACGGATTTAA
AATCCAAGAGTCTGGACAGTCTACTTCTCTCTATAACTATAAAAAATAGCT
45 ATGAAAAAATCTATCGTCATAGATTTCTTCATAGCTAATCTTAGTATGTTT
TTATTTTATTGAAATTAATTGTTTATTAGCGATATAACTTGTCAATGATAG
AGTCGTTATGTCATTATAAATGTAAGGGCTTACATATATGTTTTGCATCAA
CATTTAGTCTGTTTCATACAATTCAATTTTTCGATTATGACGATATTGATT
CATTTTATTGTTATATTAAAAAATAAGTTTGAGAGTCGATAGAAATCAGTT
50 ATGATATTGTTAATAAGTTTTTATTCGAAAACGAAGGGGAGATTTTATGAC
TCAATCTGAAAAAATTATTAACTTAACAAATCACTATGGGGCGCATAACTA
CGTTCCACTTCCAATTGTTATTTCTGAAGCAGAGGGTGTATGGGTGAAAGA
TCCGGAAGGTAACACGTATATGGATATGCTTTCGGCCTACTCGGCAGTGAA
TCAAGGTCATCGACACCCAAGAATTATTCAAGCATTGAAAGATCAAGCAGA
55 TAAAGTCACTTTAGTATCACGTGCTTTTCATAGTGATAATTTGGGTCAATG
GTATGAGAAAATATGTAAACTCGCAGGTAAAGACAAAGCATTGCCTATGAA
TACGGGAGCAGAGGCGGTTGAAACAGCTTTAAAAGCTGCTCGTCGTTGGGC
TTATGATGTTAAGGGTATTGAGCCGAACAAAGCTGAAATTATCGCTTTTAA
CATAAAAGCTATCAATAGCAATACAAATACATTGACGACCAAATTTCTCGC

TTGCTTGTCGAATTAATTCTGGATGTTTTAAAGCGCTTGAATTGAGTGATA
TTTTATCTGCTCCGTGATTTAATAGTTGCGTAATATCATCTAAATTTTGAA
TCCCTCCTCCTACTGTTAAAGGGATAAATAATTGTTTTGCCGTTGCTTCTA
TCACTTCTATCATAAGATCATGTCCTGCTTCCGTCTTCGAAATATCAAGAA
5 AGACTAGTTCATCTGCACCGGCTTCATTATAATAAAGAGCCAAATCAACTG
GATTACCGATATCTCTTAATGACTGGAACTGGATACCCTTTACGACGCGTC
CATCTTTAACATCTAAACATGGAATCACTCTTTTTTTAATCAAGATAATCC
CTCCCAGAATTCATCCAGATGTGCTGCTTTTCCTACAATAGCAGCATGAAC
ATTTAACGATTCTAATCGAAACAAGTCCTCTTGATGTCTAATACCTCCTGA
10 AGCAATTACAGGCAACGATGTATAAAGTGCGAGACGACCTGTCCTTTAGAG
CACAGTGGCGATGATATC

Sequence 3374
15 step.1002b12.cons.ok
TATTTAGCTGCTATCAAAGAACAAAATGGTGCAGCAATGAGTTTAGGACGT
ACTTCAACATTCTTAGATATTTATGCTGAACGTGATTTACAAAATGGTGAC
ATCACTGAACAAGAAGTTCAAGAAATCATTGACCACTTCATTATGAAATTG
CGTATCGTTAAATTCGCGCGTACGCCTGAATATAATGAATTATTCTCTGGA
20 GATCCAACTTGGGTAACTGAATCTATCGGTGGTGTAGGTATTGACGGACGT
CCAATGGTAACTAAAAACTCATTCCGTTTCTTACACTCATTAGATAATTTA
GGTCCAGCACCAGAACCAAACTTAACAGTGTTATGGTCTACTCGCTTACCT
GAAAACTTCAAAATCTATTGTGCTAAAATGAGTATTAAAACGAGCTCAATC
CAATATGAAAATGATGACTTAATGCGTGAAAGCTATGGCGATGATTATGGT
25 ATCGCTTGCTGTGTATCTGCCATGAAGATTGGTAAACAAATGCAATTCTTC
GGTGCACGTGCTAACTTAGCTAAAGCATTACTTTACGCTATCAATGGTGGT
AAAGATGAAAAATCTGGTAAACAAGTTGGGCCAAGTTATGAAGGTATTAAA
TCAGACGTACTAGATTATGATGAAGTCTTCGAAAGATATGAAAAAATGATG
GACTGGTTAGCTGGCGTATATATCAACTCATTAAATATCATTCACTATATG
30 CATGATAAATATAGCTATGAACGTCTTGAAATGGCTTTACATGATACAGAA
ATTATTCGCACAATGGCAACTGGTATTGCCGGATTGTCTGTAGCAGCTGAC
TCTTTATCAGCGATTAAATATGCACAAGTTAAACCTATCCGTAACGAAGAA
GGTCTTGTAACTGACTTTAAAATCGAAGGCGACTTCCCTAAATATGGTAAT
AATGACAGTCGTGTTGATGAAATTGCAGTAGATTTAGTTGAACGTTTCATG
35 ACTAAATTACGTAGCCATAAAAACATACCGTAATTCTGAACACACAATGAGT
GTATTAACAATTACTTCAAACGTTGTTTATGGTAAGAAAACTGGTAACACA
CCAGATGGACGTAAAGCTGGCGAACCATTTGCACCTGGCGCAAACCCAATG
CATGGTCGTGACCAAAAAGGTGCATTATCTTCACTAAGTTCAGTAGCTAAA
ATACCTTACGATTGCTGTAAAGATGGTATCTCAAATACATTTAGTATCGTA
40 CCGAAATCACTAGGTAAAGAAGAAGCAGATCAAAATAAAAACTTAACTAGT
ATGTTAGATGGTTATGCAATGCAACATGGTCATCACCTCAACATTAACGTA
TTTAATAGAGAAACATTAATTGATGCAATGGAACATCCAGAAGAGTATCCA
CAATTAACGATTCGTGTATCTGGATACGCTGTAAACTTCATTAAATTAACA
CGTGAACAACAATTAGATGTTATTTCACGTACATTCCACGAATCTATGTAA
45 TAAATTTAAGGTGGGGAGCAAAATGCTTAAAGGACACTTACACTCCGTTGAA
AGTATGGGCACTGTCGACGGGCCAGGACTTAGATATATATTATTTACTCAA
GGTTGTTTGTTAAGATGTTTATATTGTCATAATCCAGACACTTGGAAGATT
AACGAACCATCAAGAGAAGTGACGGTTGATGAAATGGTAAATGAAATCTTA
CCGTACAAACCTTACTTTGAAGCTTCAGGTGGTGGGGTAACAGTCAGTGGT
50 GGCGAACCATTACTACAAATGCCTTTCTTGGAGCAATTATTCAAAGAATTA
AAAGCGAATGGTGTTCACACATGCATTGATACTTCTGCGGGATGTGTGAAT
GATACACCAGCATTTAATCGTCATTTTGATGAATTGCAAAAGCATACAGAT
TTAATCTTATTAGATATTAAACATATTGATAATGATAAGCACATCAAATTA
ACAGGCAAACCTAACACACATATTTTAAAGTTTGCACGTAAATTATCTGAT
55 ATGAAACAACCTGTTTGGATTAGACATGTTTTAGTACCTGGTATTTCGGAT
GATAAAGAAGATTTGATAAAACTAGGAGAATTTATTAATTCTTTAGATAAC
GTTGAAAAGTTTGAAATCTTACCATATCATCAACTCGGTGTGCATAAGTGG
AAAAAATTTAGGCATCCCTTATCAACTCGAAAATGTTGAACCATCTGACGAT
GAAGCGGTTAAAGAAGCTTATCGCTATGTTAACTTTAATGGCAAAATACCC

GTAACATTATAGAAAGATAAATATAGTAATAGCCTGAGACATCAATCGTTG
TCTCAGGCTTCTTTACGTTTAATCAGTCGATAAACTGAGTATAAAATATAT
CTCAATTAACTTTATTTAATAAAAATATGATTCATAAGTTTAAATAGTTTA
TTTTTCTAGATATGCTTTAAATACATCTTTATAAACATCGATGTAACTTAA
ATACATATCTTTTCTAATAAATTCATCAACTTGATGCGCCATAATAGATTC
ACCAGGTCCAAAGACAGCAAAATCCACGTTATTTTCATTTGTTCCTAGGAA
ACTAGATGCATCAGTTGTACCAATCAATGCACTCACAACAATGTCTTCATG
TACATAATTCGGTGCAATACGTGTAATATTTTGAATAAGAGGATTATCACG
ATCACTTGCCACTGGATCGTGACTGCTAGGTATATCTACAGTTAAATAATC
TTCGCCCACATGACGAATGACTTTTTCAAATAAATCCTTCACGAAAGTACT
GTCATATTCTGGAACAGTTCTTACATTATATTTAGCTGTCGCTTTATGAGG
AACAGAATTAACTTGTTTACCACCATTGAATACAGAGTTTAACATTACAAA
TCCAGAGTAGATATGTGATTCTTCTTCACCAATTTTTCTGTGGAGATGTTT
CTCAATCATTGGAACAGCGTCTAACTCGTGTACTTTATCATGTTCTTTAAT
ATTTTTATATTCTTGTTTCATTTCATTTACAAAATCAACTAAAATATCAAC
AGCATTTGTACCTAGGTGTGGCATGGAACTATGTGCAGCCTTACCTTTCGC
TGTGACTACACAGCTCATAGAACCTTTATGAGCATAGTAAGCGATATTACT
AGTTGGTTCGCCAATAATTAAACCACTCACATCATCTAAGTAACCTTCATC
AGCTAGTAATTGTGCACCATATTGTTCCGTCTCTTCCTGTAGTAGCTAA
TAAACGTATCGTACCTTGCTTTAATGCGTTTGATTGTTTTAATTCAATCAT
CGCGATGACCATCGCCATAAGACCACCTTTCATGTCAGTGGTACCTCTGCC
GAATAACTTGCCATCTTTATCTGTGAGTTCAAATGGAGGAAAAGTCCAATC
ATCATGGTCACCAGCATCTACAACATCCATGTGTCCACTGATTGCTAAGAC
TGGAGCGCCACTACCAATTTCAGCAACTAAATTAGCGCGTGAATCATTAAC
TTTAACAATCTTAGAATCAATATCATATTGACTTAACAGGTCTTTTAAGTA
CTCACATACTTTAATTTCATGATCATTCTCAGTTTGAATCTTGACGATATC
AGCTAATAATCTTATTTTGTCTTGCTCACTCAATACAGTCATTTCACAGTC
CTCCTCATTTTTAGTAATGATTTTCACTTTTTAATAAAAAATAAAAATCCT
TGTACACTTAATATTATCAAAAACTTTTATTAATTAAAAATTATATGCTAA
AATAAATGGATTTTTAATAAAAAATAAATATTAACTTAACATTTCTTATAA
AAAACAACAATGTATAGAAATATTAAAAATTTTTCATTATTTATAGTTTGA
ATCGTGTTATAAAATGAAAAGCTATGA

Sequence 3375
step.1002c01.cons.ok
TCAACATTATAATATTGATTAATTAAATAACAATTGCAAGTATAATTTAAA
TATTTCTCTAATATACAGTGTCAATTTATTTTATTCACATAAGAAAATAGC
TATGAAGAAATCTATCAATTTAAATTTCTTCATAGCTAATTTTTTTCATTT
AAATTTATTGACGGCTTGAAAAATGAGTCAAAATCATCAATAACATCAAAT
TGCAAATATATTCCTTTGGTAATGGATTGACCATTAAACTTAATTCGAATT
CTATTCTTTTCTATACAATGAAACGGGTGTCATACATCATCGGTAACTAAT
TATGATAGATATGAACTTGTGGTTCTTTATCGTCTTTAGTTTTACTAATGA
GAGCACGTGGAGTATTTCCATCTTTGATTCTAATTTCATACTCATCTAGTT
TATCAAAATATTTTTCGGCTTGCTCTGTAACATATTGTGTAATACCTATCG
TTTCTGCCTGTCCGTAATAATCTATAGGCAAATCAACTGTAAGTTGTTTAG
CTTTTTTATTTACGAATTTAACCTTACCAACTGCTTGTGTGAAGTTTGAAA
AATACGATTGCAAATTATCATTAAACTGTTTAAAGTTATTATTCAACGTTT
CATCATAATCAGCTGCAGTTGACGAAGGAATTAAGGCTGCTTTTTCATTAA
TATTATCCCATGAGTTAATTTTAGTTTTACCCTCTTCAACCGTAGTACCAA
CTATAAATTCACCTGGTGTAATGGAATCTTGACTTGATTGTTTATAGATAG
CAAAATGAATAGGAATATCTTTCAAATCACTATTCTCACGTAAACGAGAAA
GCATTTCACTAGCCATCTGTTTACCTTGCTTTTCAATCTCTTTATCAGATA
AATCTTTACTAAATGTTTCGCCATCTTTCTCTTTTTTGTAATAATAAACAC
TATTCATAGCTAAACCAATTGTCATCCCTTTTATATTTTTACCTTTAGAAT
CACTATTTCCATAAAAATCCTGCTCGAGTATATTTGAAAGATAGGCTGGAG
AATTTTTAGCTATTTTCTCTTCATCTGTTTCACCATTGTGAGATGGATTGA
GTCCAAGATTCTCATTAGCATTTTTGCTTTTCTTTTCTTTTTCGCTCATCT
TGTCAATTTCTTTTTTCGTATACTTCGGATCTAAGTATGCATTAATCGTTT

TTTTATCTAAATATTGTCCATCTTGATATAAATACTTATTTGTTGGAAAGA
TTTCTTTACTTAATTCTAGTAAACCACTTTCAAAATCTTCTCCATTATAAC
CATTTGCCATATTATCTTGTAATAATCCACGAGCCTGGCTTTCTTTGAAGG
GTAATATAGTCCTATAGTTATCACCTTGAACTTTTTTATCAGTCGCTATTT
GTTTCACTTGATTTTTATTATGGTTATCCTTATGTTCACTTTGTTCTTTAT
CTGATGAAGTTTGTTTATGTCCATCTCCGCAAGCCGTTAATAATAACAGTA
TCGACATGAGTAAAAATATTGTTCGCTTCATTCACGTACTCCTCTAATTAT
TAGATTCCATTTTGTTTTTCAATAAATGCTGCTTCAGTCCAAATTTCAGTA
CCATACTTCTCAGCTTTGGCTAATTTAGACCCTGCATCTGCTCCAGCTATG
ACAATATCAGTACTTTTAGTCACGCTGCTTGTAACTTTAGCACCTTGCATT
TTCAACCATTCAGATGCTTCATTTCTCGTCATTTGCTCGAGTTTCCCTGTT
AATACAATTGTTTTCCCACTAAAATCAGGATGACCTTCGATTTCAGTTGTT
TTAATTCCTTTATAAGACATATTAACATTTTTATTACTTAATTTTTCAATT
AATGAACGAATATCACTATTTTCGAGATATGTTACAACAGATTGTGCAAGT
TTATCTCCAATATCTTGAATTTCAATTAATTCACTTTCAGTTACTTTAAAA
AGTTGATCCATCGTTTCATATCGCTCAGCAAGTACTTGACTAGCTTTTACA
CCTAAATGTCTAATACCAAGTCCAAATAATAAATGCTCTAATGACTGTTCT
TTAGATTTTTCTATCGCTAATAAAAGATTATCAACTTTCTTCTTTCCCATT
CGCTCTAATGGTAATAAATCTTCTTCTTTCAAATAGAAAATATCTGCGACA
TCTTTGATTAACTGATTTTCGTATAGCTGATGAATAATTTTAGTACCTAAA
CCATCTATATTCATCGCTTGTCTTGAAACGAAATGTATAAGTCCTTCAATA
AGCTGTGCCTGACATTTTGGATTAATACAACGTAAAGCAACTTCTCCTTCA
ATACGAACTAATTCATGTCCACAACTAGGACAATGTGTTGGCATATGATAA
ATTTCCGATTCGTTAGGTCGTCTATCTAAAATACTTTTTACAACTTCAGGG
ATGATGTCCCCGGCTTTTTTAATAACAACACTATCTCCGATACGTATATCT
CTTTCATGTATTAAATCTTCATTATGAAGTGAGGCTCTTGAAACTGTAGTA
CCAGCTACTTTTACAGGTTCTAGAATTGCAGTTGGTGTCACAACACCCGTA
CGCCCAATACTTAGCTCAATATCCAATAATTTTGTAATAACTTCTTCAGCT
GGAAATTTATAAGCAATCGCCCATCTTGGAGATTTTTGCGTATAACCCATT
TCCTCTTGTTGAGATAAATCGTTAACTTTTATAACAATACCATCAATATCG
TAAGATAAAGATCCTCTTTTGCTTGTCCATTTCTCTATATAATTAAGTACG
CCCTCAATATCTGATACTCGTTCACGTTCTTGGTTAGTTTTAAAACCTAAT
TGGTCCAATTCCTCTAGCGCTTCACTTTGTGTTGTTGCATTAAACTCGGTT
AGGTCATTCACACTATATAAGAAGACGCTTAACTTTCTTTTCGCAGCTAGT
TTAGAGTCAAGTTGTCTTAAAGAGCCTGCAGCAGCGTTTCGTGGATTTGCA
AAAGGTTGTTCACCATTTTGTTCTTTTTCATTATTCAAATGAATGAATGAA
CGACGTGGCATATAAGCTTCCCCACGGACTGAAGGATATCAGGTTACGACA
ACAACATGTGGTAAAGAAGCACTAAAATTATTATCTTCAGATATAGATATT
ATGATTTTAGACATTATGATGCCGGAAGTTAGCGG

Sequence 3376
step.1002c03.cons.ok
GGTTGTAAAAGATGTAAAAGGAATAGACGTTAGTGGCCCATTCCCACGTAT
GACATATGCAGAGGCTATGGACCGTTTTGGTTCAGATAAACCTGACACTCG
TTTCGGTATGGAACTTATCAATGTGTCACAGCTTGGTAAAGAAATGAATTT
TAAAGTTTTTAAAGATACGGTAGATAACAACGGCGAAATTAAAGCAATTGT
CGCAAAAGACGCTGCAAATAAATATACACGTAAAGACATGGATGCATTAAC
AGAGTTTGTAAATATATATGGTGCAAAAGGATTAGCTTGGGTTAAAGTTGT
TGATGATGGTTTAAGTGGCCCAATTGCTAGATTTTTCGAAGATGTTAATGT
TGAAACACTTAAACAGTTAACAGAAGCTAAACCTGGAGATTTAGTAATGTT
TGTAGCTGATAAACCTAATGTTGTTGCTCAAAGTTTAGGGGCTTTAAGAAT
TAAATTAGCAAAAGAATTAGGTTTAATTGATGAATCAAAATTAAATTTCTT
ATGGGTAACTGATTGGCCGTTATTAGAGTATGATGAAGATGCAAAACGTTA
TGTAGCAGCACATCATCCATTTACTTCACCTAAAAGAGAAGATATCGAAAA
GCTAGACACTGAACCTGAAAATGTACAAGCCAACGCTTATGATATTGTTCT
AAATGGTTATGAACTTGGTGGTGGTTCTATAAGAATACACGATGGTGAATT
GCAACAAAAAATGTTTGAAGTATTAGGATTTACTAATGAACAAGCTCAAGA
ACAATTTGGTTTCTTATTAGATGCTTTTAAATACGGTGCTCCACCTCATGG

883

```
TGGCATCGCGTTAGGTTTAGATAGACTTGTGATGTTATTAACAAATAGAAC
AAACTTGAGAGATACAATTGCATTCCCTAAAACAGCATCAGCTACATGTCT
TTTAACTGACGCTCCAGGAGAAGTATCTGATAAACAACTCCAAGAACTCTC
ACTAAGAATCAGACACTAGGGCCCCAACAAAGAGAATTTCACAAAGAAATT
CAACAGGCAGAGCAAGTTGGGACTAGGGCCTTAATTTAATGTTTCTTAAGA
AATTAGCACAGAGAGCTTAATTGTACGAATTGAATATAAAACATCGGTGTG
TAATATAATTAATCAGACATTATATTATAATTTTGGTAACTAGTTTAGAAT
ACACGAGAAAATAATACACATTTGAGTTAAATGATAAGTCTTTAAAATTTA
CTTCTCTTTTTAAATTGTGGTATGATATTTTCATAAGATAGTCATCTGTGG
TGTTCGTAAGTTTGCTTTTTATTTGGGCCTAACACTCTTTGATCAGGGAGC
CCAATAGGTTTTCTCGCAGCGCACACGCCTCTATAGGAGGACTTGCAAAAC
GAGAAACAGGGCACCCACCTGTATATAGCAGGCCGAATGATCAAGCTATTT
ATAACTACGGCATCAACGGACTCTATCGGTACGCAAGACTTTTGTCTTGCG
TATTTTTATGTATATAATTATAAAGAAGCATAGATATAATGAATTGGAATC
AAAGGAGCTAAAGATGAAACATCAATTTTCAAGGAATGAATTAGCAATAGG
ACAAGAAGGGCTGAACTTACTAAAAAATAAGACTGTTGCAGTTTTAGGTGT
TGGTGGCGTCGGGTCATTTGCAGCTGAGGCATTGGCTCGGACTAATATAGG
GCACATCATACTTATAGATAAAGATGATGTCGATATTACAAATGTGAACAG
GCAAATTCATGCACTGACTTCAACTATTGGTCAAAGTAAAGTCACGCTAAT
GGAAGAAAGAATCAAATTAATAAATCCCGATTGTAAAGTAACTTCTTTGCA
TATGTTTTATACCGAGGAAACATACAAAGATATCTTCAATAATTATGATAT
TGATTATTTTATTGATGCAAGCGATACAATCATTTATAAAGTTCATCTCAT
GAAAGAGTGTTTAGAAAGAGGAATTGAGTTAATTTCAAGTATGGGTGCAGC
AAATAAGACTGACCCGACACGTTTTGAAATTGCAGATATTTCAAAAACACA
TACTGATCCTATGGCTAAAGTAATTAGAAATCGTTTAAAACGCCTTGGTAT
TCGTAAAGGTGTTAAAGTAGTATTTTCTGATGAAAGTCCTATTGTTATTCG
CGAGGACGTAAAAGAAACAGTAGGAGATAAAAATGCAATCAATAGAAAAGG
GCAAATGCCTCCATCTTCTAATGCATTTGTTCCAAGTGTAGTAGGCCTTAT
TTGTGCAAGCTACGTTGTCAACGATATTTTAAAAGATATACCTGTAAGGCG
AATTAAAGATAAAGGACAAAATTAAAGTTATCTATCTCATTGCAAGAAAAT
CATCATCTTCTATTTTTGTGAAATTTGATTCATAGTTTAAAAAAAACAAAA
TGTATTTTAAATAAATCTACAACAACTCATAGTAAAAAAACCTGAGGCATC
AGTATTGTCCCAGGTTTTTTGCATCTAAGATATTAATTTCAACTACATATT
TTCAATGAATAAAATTATGTTTTATTTCGTAAGTTATCGTAAATTGTTTTA
AATTGTTGTTCACTTTTTGATGTGGTTTTTGGTTCATAATACACTCTGTTT
TTTAATTTTTCTGGTAAGTACTGTTGTACAACATGACCATTTTCATAGTTA
TGAGGATATTTGTAACCTATAGCCCGTCCTAACTTTTTAGCACCTGAATAA
TGACCGTCTTTCAGATAATCCGGTATTTGACCTATATGTCCATTTCTTATA
TCGCCAAGAGCTTTATCTATTGCTGTGATACCAGAATTTGATTTAGGTGAT
AAGCATAATTCTATTACAGCTTGACTAAGGGGAATTCGTGCTTCTGGAAAA
CCTAGACGTTCAGCTGATTGTATTGCTGCTAAAGTCCTCTGACCAGCATTA
GGTGATGCTAACCCAACATCTTCATAACTAATTACAAGTAATCGACGTACG
ATTGTAGGTAAGTCACCAGCTTCAATTAGACGTGCTAAGTAGTGGAGTGCT
GCATTGACATCACTACCTCTAATTGATTTTTGAAAAGCGCTCATGACGTCA
TAATGCATGTCTCCATCTTTATCACTTACAAAAGCACCTTTTTGTAAGCAA
TCTTTTGCGTCATCTAATGTAATATGTCTTTCGTTTTCTTCACCAATATGA
GCACTTAGTACAGCTAATTCCAAAGCGTTTAATGCACTCCTAACGTCGCCT
TGGCTTTGCGTTGAAAAATATTCGATAGCATCTTCATCTACAATAGGATGA
TAAGTACTTAAACCTCTTTCCTTATCATTTATAGCTCTATCTAATGCCAAT
CTTATATCATCTTGATCTAATGGGTAAAGCTCAAATATTTGTGCTCTAGAT
CGTATCGCTGGATTTATCGCATGATAAGGATTTGAAGTTGTTGCACCTATT
AATACAATTTTGCCATTTTCAAGATGTGGTAATAAAAAATCTTGTTTAGCT
TTGTCTAAACGA
```

Sequence 3377
step.1002c04.cons.ok
ACATTATTGCGTTTAATTGCAGGTATTGAGAATGCAGACGAAGGACGTATT
CAATACTTCAATCAATATTTGTCAAGACGTCGAATACGTCATATTGTAGGC

```
          TATGTCCCTCAAGACATCGCACTATTCGAGCATATGACTGTCATGGAGAAC
          ATTGAGTTTTTCAAGTCACTTTGTGAAAATCCTATTTCAGATGAAACACTT
          CATTCTTATTTATCACAATTAAATTTTACTGATACAAAAGTGAAAGTATCT
          AACCTTTCTGGGGGAAATAAACGTAAAGTCAATATTATGATAGGTCTACTT
     5    AGTCGGCCTAAAATACTTATTCTAGATGAGCCAACAGAAGGCATTGATTTA
          GAATCAAGATATGATATTCACAACTTATTACAACAAATGACCGATCAATGT
          TTAATCATCATGACGACACATCATTTAGACGAAGTTGAAGCACTAGCAGAT
          GATATTAAAGTTATAGGTCAAAATCCTTTTTATCATGATATTTTAGAAAAT
          AAAGGTTGGTCTTTTAAAAAATATGCAAATGCCTTAGCTGATAATACGAAA
     10   TCTTAAGCTGTACTTTCACTCCATCACTTCATCAATATTTAACATCAAAAA
          GGCTTGCCCTTTCGTCGTATATTTCTCGATTGAAAAGGCAAGCCTTAATAA
          ATTTTATACACATTTATTTATCGTGTCTAACTTATCTGTACACTAATCCAC
          CATCAGTTAAAATGGATTGACCTGTAATATAATCAGAATCATTTGAAGCTA
          AGAATGATACTAAGTTCGCAACATCTGATGGCTCTTGATATCTTCCAAGTT
     15   TGATTTCTGAAGAAAATGCTTCAAACGCGTCTCCTATTTCTAGACTATCAT
          CAAGTTTCACCATTTCTTCATCAATACGATCCCACATTTCAGTTTTGGCAA
          CTCCTGGGCAATAAGCGTTTACGGTAATGCCTTTGTCTGCTAATTCTTTAG
          CAGCTGTTTGAGTGAAGGAACGTACAGAGTGTTTTGTAGCTGAGTAAGTGC
          CGAGTACTTCATAAGATTCATGTCCTGCAATACTACATGCGTTGATAATTT
     20   TACCTTTACTTTTTTGTTTAATAAATTGGTTAGCAGCTGCTTGAATTCCAA
          ACAATGTACCAAATACGTTAATATTAAATAACTTAGATAATTCCTCTTCAC
          CAATTTCTAAAATTGGCGTCACTGCATCAACACCTGCGTTATTCACCATGA
          CATCTAATTGACCAAATTCTGTAACTGCGAACTGAACTAATTCTTCCTGTT
          CTTTCTTTTTAGATACATCACTCTTGAATGCGACAGCTTGATAACCTTTTT
     25   CTTTGAATTCTTTCTCAGTTTCTAATAGAAGTGCTTCATTAATATCTTGTA
          GTACTATATTAAACCCATCGTTAGCAAGACGTTCTGCAATACCTTTACCTA
          ATCCACCTGCTGAACCTGTTATAATTGCTGTTTTACTCATATCTATCTTCC
          TCCTTATTTATCATCTACCCTTTAAACATGGTGGTAGTAATAAGATTGGGA
          AATGTCAAATGTTGTTAATATATTTGTCGAAAAGTTGTTGTTTTTATGCAT
     30   TATTAAATACAACTTCAAATTACTAACACTGCTTTTTACTGCAACAATATA
          CCTTTTAGATATCTACATTCTTTTTAGGAATTAAACAATACGCCGCAATAA
          ATGCGATAAACACTAAACCTGCACCCACGATAAATGCGATAGTAGCTGCTA
          AATATGCCGGGTGATTTACAGATAATGCCGTATATATTGTTGTCACAACAG
          CGATTCCAAATGCTGCACCTAGTGTAGAAGTCATTTTGATAATTCCTGATG
     35   CAGTACCTGCTTTTTCAGCTGGAACATTAGATACAGCTGTAGATAGCGCGG
          GTGTAGCAAAAAATCCTAATCCTAGACCTATAAAGATAAAGCCAATACATG
          CCACAATATAATAAATCATATTAGGCAAAGAAGTGAATGCTAAAAGTATAA
          TACCAACAGTAATTGATACCGGACCTAACATCAATGGCAATTGCGGTCCTC
          TTTTTTGCATAAAACGTTCACCAACACGAATCATTAACAAACTACACAGCA
     40   TATATGGAATTGTAATTAAACCGGCTTGTGCAGCTGATAGATGTTTATCGT
          CTTGAACATAAATATTAAATAATGCTAATGAACCAATATCCATGTTCACCA
          TCAAGTTGGCTAATGTTGTTCCAATATAAACATTATTTGAAAATAAACTTA
          AATCTATAAAAGGTTCATCTTGACGTTTTTCAAATATATAAAATGCAATTA
          ATGTCACAATGAATATAGCAATTAATATCAATATAAGAGGATTTAACCATC
     45   CAATTCTATCTCCTTGCGTAATCACTACGTTAATGCTTAACATCATGACAA
          CAAAAATAATGATTCCAACGATATCAAATTTTTTATTATATGCAGACTCAT
          CTTTAGATTCTGGTATACCTCTAAGCAATATGAGTGAATTAAGTTTAAATA
          ATATAATAAGAGAATTTAACAAGCGTTTTGAAGAAAAATATAGTCAAAGTT
          ATGAATATAAAAAATTAGATTTTTCTGCTACTACAACCATTTATGATTCAG
     50   AAATATCAGAGTTTAAAGATTGGGTAGATGCAAATTATTTAGGTACTAACG
          TTGAAAATAACATTCAAACTGAAAAAAGATTTTTATATGAAAGACCACCAG
          TTAGATATGATAGTGTAACACCTGAGTTAGAGTTGTTAAAAAGAAATTACG
          ATAAAAATGTAACTGTATTTGGTAATTTGCCTAAAAAAGCGATACAAGTTC
          CTAAATATACTGGTGGCACTACTACGCCTGATTTTGTCTATATGATAGAAA
     55   CTGATGAACAAGATGCAAAATACCTTATTGTTGAAACAAAAGCAGAAAACA
          TGAGACTAGGAGATAAAAGTATTGGTGAAATACAAAAAAAATTCTTTAACA
          CATTAGATAATTTGAATATTAAATATCAATTAGCTACTAGCGCGCAAGATG
          TTTATAATGAAATTAAAAAAATTAGATGATTCAAAGTGAGGGAATATGATGG
          GGAAATCAGAAAAAATTTCATTACTTGAAAAAGTCCAAGATGGTTTAGTAG
```

					885

ATAAAACCGGAGCAAAACCTAAGTTACCAACAACTATTACCGATTTAAATC
AAGAAACTTTAGAGGTATATAGAATACCACTAAAGTTTCTGTATTATAATG
ATAGAAATGGAAGAATTGCTTCTGTAATATCCAGAGTAAGCGACGATATAA
AAGTTGCTTATGAATTTGAAGATAATAATTATAATAAAAGTATTGAAAATA
5   TGATATATGAGGCTAATACTTCTGCTTTAAAAAACACTAAAAAATCTATTA
AAGATAAA

     Sequence 3378
10   step.1002c05.cons.ok
     AGATATCTCACCATTAGGTGCTGCAGCTTTAAGTGGAACTACACACCCTAT
     AGATAGACATCTGACTCAAGAACTTCTAGGATTTGCAAATCTGTATGAAAA
     TAGTTTAGATGCTGTCAGTGATCGTGACTATATCGTTGAAACATTGCATCA
     TATTTCACTCACTATGGTACATCTATCACGTTTTGCAGAAGAAATCATTTT
15   TTGGTCGACTGATGAGGCAAAATTTATCACTTTATCAGATGCGTTTTCTAC
     TGGCTCATCTATTATGCCACAAAAGAAAAACCCTGATATGGCTGAACTAAT
     AAGAGGAAAAGTCGGACGTACTACAGGACACTTGATGAGTATGTTAGTAAC
     ACTTAAAGGCTTACCCTTAGCTTATAATAAAGACATGCAAGAAGATAAAGA
     AGGTTTATTTGATGCTGTACACACACTTAAAGGCTCTCTTCGAATCTTCGA
20   AGGTATGGTTGCATCTATGAAAGTTAATTCAAACCGTTTAAGTCAAACAGT
     AAAAAATGATTTTTCAAATGCAACAGAATTAGCAGACTATTTAGTCAGTAA
     AAGTGTACCTTTTAGAACCGCTCATGAAATCGTTGGTAAAATCGTATTAAA
     TTGTATTCATAAAGGTATATACCTATTAGACGTACCTTTAAGCGAATATCA
     AGAACATCATGAGAATATTGAGGAAGATATATATGATTATTTAACACCTGA
25   AAATTGTCTCAAGCGTCGCCAAAGCTATGGTTCAACTGGTCAAGAATCAGT
     AAAACATCAACTAAAAGTCGCAAAAGCATTATTAAAAGACAACGAATCAAA
     ATAGTTATTAAAAATAATAGCCATCTAAAATTACAAAAATTTAATTTTGTA
     TTAAGTTATCAAAGTTAGTGCTAAGTGTTAGGGGGTTTCCGCCCCTTAGTG
     CTGCAGCTAACGCATTAAGCACTCCGCCTGGGGAGTACGACCGCAAGGTTG
30   AAACTCAAAGGAATTGACGGGGACCCGCACAAGCGGTGGAGCATGTGGTTT
     AATTCGAAGCAACGCGAAGAACCTTACCAAATCTTGACATCCTCTGACCCC
     TCTAGAGATAGAGTTTTCCCCTTCGGGGGACAGAGTGACAGGTGGTGCATG
     GTTGTCGTCAGCTCGTGTCGTGAGATGTTGGGTTAAGTCCCGCAACGAGCG
     CAACCCTTAAGCTTAGTTGCCATCATTAAGTTGGGCACTCTAAGTTGACTG
35   CCGGTGACAAACCGGAGGAAGGTGGGGATGACGTCAAATCATCATGCCCCT
     TATGATTTGGGCTACACACGTGCTACAATGGACAATACAAAGGGCAGCGAA
     ACTGCGAGGTCAAGCAAATCCCATAAAGTTGTTCTCAGTTCGGATTGTAGT
     CTGCAACTCGACTATATGAAGCTGGAATCGCTAGTAATCGTAGATCAGCAT
     GCTACGGTGAATACGTTCCCGGGTCTTGTACACACCGCCCGTCACACCACG
40   AGAGTTTGTAACACCCGAAGCCGGTGGAGTAACCATTTGGAGCTAGCCGTC
     GAAGGTGGGACAAATGATTGGGGTGAAGTCGTAACAAGGTAGCCGTATCGG
     AAGGTGCGGCTGGATCACCTCCTTTCTAAGGATATATTCGGAACATCTTCT
     ACGAAGATGAGGGAATAACGTGACATATTGTATTCAGTTTTGAATGTTTAT
     TAACATTCATTTGTACATTGAAAACTAGATAAGCAAGTAAGATTTTACCAA
45   GCAAAACCGAGTGAATAGAGTTTTAAATAAGCTTGAATTCATAAATAATCG
     CTAGTGTTCGAAAGAACACTCACAAGATTAATAACTAGTTTTAGCTATTTA
     TTTTGAATAACAATTCAAAATATGGTGGAAACATAGATTAAGTTATTAAGG
     GCGCACGGTGGATGCCTTGGCACTAGAAGCCGATGAAGGACGTTACTAACG
     ACGATATGCTTTGGGTAGCTGTAAGTAAGCGTTGATCCAGAGATTTCCGAA
50   TGGGGGAACCCAGCATGAGTTATGTCATGTTATCGACATGTGAATTTATAG
     CATGTCAGAAGGCAGACCCGGAGAACTGAAACATCTTAGTACCCGGAGGAA
     GAGAAAGAAAAATCGATTCCCTGAGTAGCGGCGAGCGAAACGGGAAGAGCC
     CAAACCAACAAGCTTGCTTGTTGGGGTTGTAGGACACTCTATACGGAGTTA
     CAAAAGAACATGTTAGACGAATCATCTGGAAAGATGAATCAAAGAAGGTAA
55   TAATCCTGTAGTCGAAAACATATTCTCTCTTGAGTGGATCCTGAGTACGAC
     GGAGCACGTGAAATTCCGTCGGAATCTGGGAGGACCATCTCCTAAGGCTAA
     ATACTCTCTAGTGACCGATAGTGAACCAGTACCGTGAGGGAAAGGTGAAAA
     GTACCCCGGAAGGGGAGTGAAAGAGAACTTGAAACCGTGTGCTTACAAGTA
     GTCAGAGCCCGTTAATGGGTGATGGCGTGCCTTTTGTAGAATGAACCGGCG

```
       AGTTACGATCTGATGCAAGGTTAAGCAGCAAATGTGGAGCCGCAGCGAAAG
       CGAGTCTGAATAGGGCGTTGAGTATTTGGTCGTAGACCCGAAACCAGGTGA
       TCTACCCTTGGTCAGGTTGAAGTTCAGGTAACACTGAATGGAGGACCGAAC
       CGACTTACGTTGAAAAGTGAGCGGATGAACTGAGGGTAGCGGAGAAATTCC
   5   AATCGAACTTGGAGATAGCTGGTTCTCTCCGAAATAGCTTTAGGGCTAGCC
       TCAAGTGATGATTATTGGAGGTAGAGCACTGTTTGGACGAGGGGCCCCTCT
       CGGGTTACCGAATTCAGACAAACTCCGAATGCCAATTAATTTAACTTGGGA
       GTCAGAACATGGGTGATAAGGTCCGTGTTCGAAAGGGAAACAGCCCAGACC
       ACCAGCTAAGGTCCCAAAATATATGTTAAGTGGAAAAGGATGTGGCGTTGC
  10   CCAGACAACTAGGATGTTGGCTTAGAAGCAGCCATCATTTAAAGAGTGCGT
       AATAGCTCACTAGTCGAGTGACACTGCGCCGAAAATGTACCGGGGCTAAAC
       ATATTACCGAAGCTGTGGATTGTCCTTTGGACAATGGTAGGAGAGCGTTCT
       AAGGGTGTCGAAGCATGATCGCAAGGACATGTGGAGCGCTTAGAAGTGAGA
       ATGCCGGTGTGAGTAGCGAAAGACGGGTGAGAATCCCGTCCACCGATTGAC
  15   TAAGGTTTCCAGAGGAAGGCTCGTCCGCTCTGGGTTAGTCGGGTCCTAAGC
       TGAGGCCGACAGGCGTAGGCGATGGATAACAGGTTGATATTCCTGTACCAC
       CTAGTATCGTTTTAATCGATGGGGGGACGCAGTAGGATAGGCGAAGCGTGC
       TGTTGGAGTGCACGTCCAAGCAGTAAGGCTGAGTGTTAGGCAAATCCGGCA
       CTCATAAGGCTGAGCTGTGATGGGGAGAGGAAATTGTTTCCTCGAGTCGTT
  20   GATTTCACACTGCCGAGAAAAGCCTCTAGATAGATAACAGGTGCCCGTACC
       GCAAACCGACACAGGTAGTCAAGATGAGAATTCTAAGGTGAGCGAGCGAAC
       TCTCGTTAAGGAACTCGGCAAAATGACCCCGTAACTTCGGGAGAAGGGGTG
       CTCTTTAGGGTTCACGCCCAGAAGAGCCGCAGTGAATAGGCCCAAGCGACT
       GTTTATCAAAAACACAGGTCTCTGCTAAACCGTAAGGTGAT
  25


       Sequence 3379
       step.1002c06.cons.ok
       TATCTAAAACCTCTTCATAAAGTGGACAAGAAGGTAAAGTTAAGTTATCCA
  30   TTTGAACTTTAATTAGATGTAATATTCTGTCTGCATCTTTATTCAATTGGT
       CATATGCCGCACTATTTAACTTCGATTGGTTAGACATGACATATTCCCCCT
       ATCGTTTCAATTAATATCTTCACAATAGAAGTTTATCCTACTGAATGTTAA
       AAAGCAAGCGAATTTAAATTAAGTGAAAGCGATTCTATGTTAAAATAATGA
       AAGATAAAAATTTAGGGAGTGGGCATTTATGATTCAAATAAAAGGAGCAGT
  35   CAATTTCCCTATTTCATTGGATAGTACGACTTGGATATTTGACGATAGAAA
       AGTTACTATTGATGATTTAGAACGTGGGGTATTTGATGGTACTAGACCCAT
       CAACTTTGATGATAACAAGGAATGGAACCGTGCTATTTTAGAAGGACAAAC
       AAATCCACCAACGCTTGATTCTGAAATTAAATATAAAAAACGATCGGTTTT
       AGATGAAACATTTGTGAATATTTATTTTTCTTTTTATCTTCATAACGTTCT
  40   TTAGCTAAATCAATCAATTTAGCAATTAAATCAGGGTAAGATAAGCCCATA
       TTTTTCCATAAGTTTGGATACATACTATATGCAGTAAATCCTGGCATGGCA
       TTTGTTTCATTAATATAAATTTGATTATCATCAGTAACAAAGAAATCTGCA
       CGAACTAATCCCGAACAATCAGTAGCTTTAAAGGCCTCTAATGCCATGTTT
       CTTAATGTCATTTGAACATCTTGATCTAAATCTGCTGGAATATCTAATCTA
  45   ATCTTACCGTCTTTATACTTTGATTTATAATCATAAAACGCTACATCCTTA
       ACAACCTCACCAGGCCATGTCGTTTCAGGATAATCGTTACCTAAGACAGCT
       ACTTCGATCTCTCTAGCATTAATCCCTTGTTCAATGACAAGTTTACGATCA
       AATTGGAATGCTTCAGCTATCCCAGATTTTAATTCTTCTTCATTGTTACAT
       TTACTTATACCAACACTTGAACCGAGATTAGCAGGTTTTACAAATACCGGA
  50   TATGTTAACTTATCATTAACTAATTTAATGATATTATTTTCATATTTTTCA
       TACTCACTTCTTAAAAAGCTAATATAAGGTAATTGAGGTAAACCTCTATGC
       TCAAATAATTGTTTCATCACGAGTTTATCCATTGAGCTTGAAGCAGCTAAC
       ACACCATTACCTACATATGGTATATCAAGTACTTCAAAAAGACCTTGGATA
       GTTCCATCTTCTCCATTTGGACCATGCAATAATGGGAATACTGCATCATAT
  55   GATTTTCCTAAACTACCTTTACTGAGTAACTGTGAGATTTCTCCAGTCTCT
       ACATCGTTAATGACGAGTTCATCAGTATTTTTTATTTCTTGTGTAATATTA
       TCTTTTTTTCTTCCATTCACCATCGTTTGTTATATAAATGATATCAACTTGA
       TATCGTTCTTTATCAATTGCGTTTAAAACATTTTGTGCAGTTAAAATTGAA
       ACAT
```

Sequence 3380
step.1002c07.cons.ok

5   TATATCGCTTTTGTTGCTTAAGGTTTAAATCGTAACCATCAACATTAATGG
ATAACGTGTAATATTTTTCTGGTTTAAAATAGTTAATCTCATCTTGTCTTG
ATTTAACTATTTGTATTGTTGGAGTTTGTACACGACCTAATGATAATTGTG
CATCATATTTTGTCGTTAGTGCACGTGTTGCATTAATCCCTACTATCCAAT
CTGCTTCACTTCGTGCAAGTGCTGCTTCATATAAATTTTGATACGCGTTTC

10   CATTTTTTAACTGTTTAAATCCTTCTTGTATGGCTTTTTTTGTAACCGAAC
TAATCCACAAACGCTTGATTGGTTTTTTTATTACCTACTTTATCTAAAATAA
GACGAGCTACTAGTTCACCTTCTCGTCCAGCATCTGTTGCTATAATAATTT
CTTTAACATTTTTATCTAAAATTAAAGATTTTACAATTTTAAATTGTCTAT
TTGTTTTACTAATCACTACTGTTTTCATTTTCTTAGGAATGATAGGTAAGT

15   CTTCTAAATTCCATTCTTTATATGAAGGGTTATATTGTTCAGGCGTTGCAT
TTGTCACAAGATGCCCCAATGCCCAAGTTACTATATACTGTTTCCCTTCTA
TATAACCATTACTTTTTTGTTGAAGATTCAAAGCATTAGCGATATCTCTTC
CGACAGATGGTTTTTCAGCTAAGATTAAAGATTTCATTTTAATTTTCCTTT
CACATTTATAAAATACGTCATGTTATAGAATTACATACATTGTAACATGTA

20   TCTTTTTGTAAATTACTATATATAAAATTAATTAAGATAGTTATATATACA
TATCTTGTTCAATCATATAGTTCAAAAACACATTATCACTTAATACGAATA
TTTACTTACTAAAAATTTCAACACTTTTAAGGTGTTGTTTTTTACTTCTTT
TTTTAATTACTATAGAATGATTTCGAGTGTAATTGGGACACTTAATTTTGT
TCTTAGGAGAGTGAGTAATAAGTGCAAATTGTAGATTTTTTGATAGCACTT

25   TTACCAGCCCTATTTTGGGGTAGTGTAGTCATTATAAATGTTTTTGTAGGT
GGTGGACCATATAATCAGATTCGAGGTACAACTTTAGGTACACTTTTTATC
GGATTTTCTCTACTTGCTACTGGACACGCAGCGTTTGATAACCTTACAGTA
ATTATTGTCGGTTTAGTATCAGGAGCTCTATGGGCTTTTGGTCAAGGTAAT
CAATTAAAATCAGTGCATTTAATAGGTGTATCTAAAACGATGCCTATTTCA

30   ACGGGTATGCAACTTGTCGGTACCACTCTATTTAGCGCTATTTTCTTAGGT
GAATGGAGCACGATTGTTCAAGTAGTGATGGGACTTATAGCAATGATCTTA
TTGGTTGTAGGTATTTCTTTAACATCACTTAAAGCCAAAAGCGAAGGCAAA
TCCGATAACCCAGAATTTAAAAAAAGCAATGGGAATATTACTTCTATCAACA
ATCGGTTACGTAGGTTATGTCGTTCTTGGAGATATTTTTGGAGTAAGTGGT

35   ACAGATGCTCTCTTCTTCCAATCAATTGGTATGGCAATTGGAGGATTAATC
CTTTCAATGAATCATAATACTTCAATTAAATCTACTGCTCTAAATCTTATA
CCAGGTGTTATCTGGGGTATCGGTAACTTATTTATGTTCTATTCACAACCT
AAAGTTGGTGTAGCAACTAGTTTCTCATTATCACAACTGCTTGTTATTGTT
TCAACTTTAGGGGGTATCTTTATTCTAGGGGAGAAAAAAGATCGTCGCCAA

40   ATGATTGGTATTTGGTCAGGTATTATCGTTATAGTTATAGCTTCAATCATT
TTAGGCAACTTAAAATAGAATTTTAAACTTTAGGGAGGTAACATAGTGTTT
GAAGAATTAGAAAATAAAGTGGTTCTTATTACTGGAGCTGCCACTGGAATT
GGCAAATCTATTGCGGAAAATTTTGGTAAAGCTAAGGCCAAGGTTGTTATA
AATTACCGTTCTGATCGACATCATGATGAAATTGAGGAAATTAAACAAACT

45   GTTGCTAAATTTGGTGGTCAAACATTGGTGGTTCAAGGTGATGTTTCAATT
GAAGAAGATATTAAACGAATGATTGAAACAACAATTAATCACTTTGGAACT
TTAGACATTATAATTAATAATGCTGGATTCGAAAATTCAATCCCAACTCAT
GAAATGTCGATTGACGACTGGCAAAAAGTTATTGACATAAACTTAACTGGC
GCCTTTGTGGGTTCAAGAGAAGCCATCAATCAATTTTTAAAGGAAAACAAG

50   AAAGGTACTATTATTAACATTTCGAGTGTTCATGACACTATTCCATGGCCT
AATTATGTACACTATGCCGCAAGTAAAGGTGGCTTAAAATTAATGATGGAA
ACAATGTCAATGGAATATGCCCAATACGGTATTCGTATTAATAATATATCT
CCTGGGGCAATTGTTACTGAACACACTGAAGAAAAATTTTCTGACCCAACG
ACGCGTGAAGAAACAATAAAAATGATACCTGCACGTGAAATTGGAAATGCT

55   CAAGATGTAGCTAATGCAGTACTATTCCTATCTTCAGATCTTGCAAGTTAT
ATACACGGTACAACATTGTACGTTGATGGTGGCATGATGAACTATCCAGCA
TTTATGGGTGGTAAAGGTTAAGTAGCATTTTTCAATTTTAACTAGAAATAA
AATGTCATAAATTAGTAATGAATAGTAGTCTTAAAAAGTATCAGTCATCGA
CTGAATATAAGTGACAAACTATAATAAAAAAGCTATGAAGATATCTACAAT

GATACATTTCTTCATAGCTTTTTTAGTTATATAAATAAGTATATTGATTAT
GCAATCAATTTTTATACAAATACAGCTTGATAAATATACAAAATGATAATG
CTAATATGAATATTTAGCAAACAATTCCGTTTCTAATGATGATCAATTATC
ATCTGAAGAACAGCAGAAATACAATGAAAATACTAACGAAAACAAACAGAA
5 TAGTCATGTTGACTTAAATTCATTACCTGCTACAGATTTTCGAACAGATTG
GATGTCTGAATCTGGTCAAAAACAAGTGAATGATTTAGCAAACCAAAAAGA
TAGCGGTCAAATTTCACAATACGAATTTAATGATCGCGTATCTGATGTGAT
GAATAATGAAATAGATAATCAACAGCCATAATTGAAATATAGGGTAGCACG
TCTACCCTTTTTATTTAGGGAGGGTAAACACGTGGCATAAGAAATTCACCT
10 ATAAACATGGTGAAACTAAAAATCGCTATTATGAGAAGTACAAAGACCCTC
TCACAAACAAATGACGACGTGTTAGCGTGGTACTTAATAAAGAATGGTAAA
CAGTCACAAAAAGAGGCTCAGAAGCGCTTAAATGAGCTTATAGAGGCAAAA
TTAAACGATAAGACACCTACTACACTTAAGTCACTAACATTCCATGATCAT
TTGATGAGTGGTTTGAAAGATACAAATTAACTTCCGGATTTAAACAATCTA
15 CCGTTACAACGAAAAGTTATAAAATAGCGCACATCAAAAGAAATATTGATA
AAAATATACTCGCTCAAAATACGAATGCTGATGTTTTACAAGACTTAATTA
ATTCTTCATTAACAGAGGGATTAAGTCATAAAGTGGTTAAAGATGATTTGA
GTATTATAAAAAATATACTACGCTATACTCAAAAGAAATATAACATTACTG
ATAAATCTTACATAGATGATGTAATAGTACCTAAAAAAGCGACTACAATAG
20 AAGAAGTTAAAGCAAACGTGAGAACTATTTAGAAATGAATGAAGTTCATC
TTATCGCTGAAGAGTT

Sequence 3381
25 step.1002c12.cons.ok
GCTTTTCTCAAATTATCTTTAACACGTTCCTCATTGACTTTAGTTTGATGT
TGTTCTCTTTTAGCTTGATTTTCCTTCTCTATTTTCTCACGTCTTATGCGC
TCTTCTCGCTCTACGTTATGAAAGTATTCACGTCTTTTACGTCGATATTGA
CTGCGTGGTATGACTTGTTTTTTATTGTTATCCACTAGAAGACCACTCCTA
30 TAGTTTCTTTAACTCATTCTTTATCTGTAAAATTCAAGTTATTTTAATACT
ATATATCATCATCATTTTAAATTCGTTTATAATCTAAAGTTTAATATTTTA
ACTATAAAACTGCAATGTTCTAATAAAAATTAAAATAAAAAGAACCTCCAC
TTAAATTATTATTAAGCAAAGGTTCTAGCATATCAATCAAATTATTTTTGA
TTACCTTTTTTCTTGTTGCCAATTGCTGAAGTTAACCATCCTATTAAAACT
35 AAGCCAACTAATACTACCCAGAAGATTGTTTGCCATAAAGCACTATGTGGA
AATGCTTCTGGTAAAACGCCAATATCAGGATGTGCAAGTACCATTATAACA
AGTTTAATACCTACCCAACCTACAATTGCAAACGCAGCACCTTCAAGTCCT
GGATATTTATTCAACAATTCTACAAACCAAGTTGCTGCAAATCTCATCAAG
ATGACACCTATCATTCCACCAAGGAACATAACAATAAATTGGCCTAAGTCC
40 ATACCACCAAAATGTATGCCAACTTTTGGTAATGTAACGGCTATGGCTAAT
GCGGCAAGCATCGAATCAATTGCAAACGCGATATCAGCGAATTCAACTTTA
AATACTGTTCCCCAAAAAGATTTAGGGCCTACTTCTTTTTCGTTGCCTGTT
TCATCGAAATGATGTTCGTCTCCTGTTTCTTTATGGTGTTTTTCATTTGAC
TGATGGAAAAATTGCCATAAATTTTTAATAGACATATAGATTAAGTAAACA
45 GCACCTGCTGCTTGAATCCACCAGAAGTTTGCAATAATACTTATTAAAAAT
AAAGCAATAAATCTAAAAATGAATGCACCTAATAGGCCATAAAAAAGTGCT
TTTTTACGTTGTTTAGGTGGTAGATGTTTAACCATTACCGCCATTACAATT
GCATTATCAGCTGCTAATAATCCTTCTAAAAATACAAGTACTAAAAGTACC
CATAAATAAGGTAAAATCAAACTCGGATCCATTGTGGACAACTCCTTTTAG
50 TTTAAATACAAATAAAGAGACCTATGCTAATTTAATAGCAAAGGTCTCACT
TAGCAATATTTTTGCCCACTTTACCGGAAGACTTATCTTCGTAATGACGAC
AAAGTGTTGAATAGTTATTATCGGTTTAAATGCATTCAATAACTATTCTCA
AGTTACTCCCCTCTGACAGTTTTGTCATAGGTATTTATTTGTTATTTTCAG
TATACATACTATGACACTGCATAGCAATATAAACATTTTTCATATACTATG
55 GTGTGCCCGAATAGTGATATGTTTAAACATTCAATTATAATAAACTATACA
GTTTAATATCAGAGAATTTTTCTTCAAACCAACGCGTTGCAAATTCATTTT
CAAATAGGAATACATAATTGTCGTATCTATCTTTTACTAAAATTGATCTAG
AAGTATTCATATTATCTTTAATATCTTCTTCATTTTCAATCCAACGCGCAA
TTTTACGTCCAACTGGTTCCATAACCACATCTACATTATATTCATTATTCA

```
TGCGGTGTTCAAACACCTCAAATTGTAATTGACCCACAGCACCTAAGATAA
TTTGATTTGTGTGCAACGTTTTATAATATTGAATTGCACCTTCTTGAACAA
GCTGTTCAATCCCTTTATGAAAATGCTTTTGTTTCATTACATTTTTGGCAG
AAACCTTCATAAATATTTCTGGAGTAAATTGTGGTAACTCTTGGAAACTAT
```
5   ```
ATTTCTGCTTACCACCAACTAAAGTATCGCCTATTTGATAGTTACCTGTAT
CATACAAACCAATAATATCGCCTGCTACAGCATGATTAACAGTTTCTTTAT
CATCGGCCATAAATGATGTGGAACGAGTTATTTTTTGTTTTTTATTAGTTC
TTTGTAAGGTTACATCCATTCCTCTTTCAAACGCACCACTTACCACGCGCA
TAAATGCAATTCTGTCACGATGCTTTGGATCCATATTGGCTTGAATTTTAA
```
10   ```
AGATAAACCCAGAAAAATCTGTGTCGAATGGACTAACATCAACCTCTTCTT
TAGTTTGTCTCGCATTAGGCATTGGCGCATGATCAACGTATGCATTTAAGA
AATTTTGTACGCCGAAGTTTGCTAATGCCGATCCAAAGAAAACAGGTGTTA
ATTCACCGTTCAGTAACGCTTCATTATCAAATGCTTCACCCGCTTCATCGA
CAAGCATCATTTCTTCAATTGCTTGTTCAAATGCACTATCATTTTTTATCG
```
15   ```
CGTGTTCTTCTTTAAGTTCGTAATCTTCGTTTAAATGTAATAAATTTTCTT
CATCTCTAAATGGTTCAATCGTTTTAGAGTGTCTATCAATAATACCAAAGA
AATTTTGTCCCATACCAACAGGCCAATTCATTGGATAAGTGTCAATATTAA
GTGTTTCTTCTATTTCATCAAGTAATTCAAATGGCTCTTTACCTACTCTAT
CCAATTTATTGATGAATGTAAAAATAGGGATACCTCGCATTTTACAAACTT
```
20   ```
TAAATAATTTTAATGTTTGTGGTTCTATACCTTTTGCACAGTCAATAACCA
TAACTGCACTATCTACTGCCATTAATGTACGATACGTATCCTCTGAAAAAT
CTTCATGCCCGGGCGTATCTAAAATATTTATCTTATAATCATCATAATCAA
ATTGCATTACAGAGCTTGTAACTGAGATACCACGTTCTTGTTCTACCTTCA
TCCAATCACTTGTTGCAAATTTACCCGTCTTTTTCCCTTTAACTGTCCCAG
```
25   ```
CTTCACGAATCGCTCCACTGAAATACAACAATTTTTCAGTTAAAGTTGTTT
TACCAGCATCAGGGTGTGAAATGATGGCAAATGTCTTTCTTGACTCTATTT
CTTCTTTTAAATTCATTTATCATTCTCCTTTGGATTCATTTCTAATCAAAT
CAGTTATATTACGTGCCAATGCATCTGCTTCAACTTCAGTCAACAAATTGA
ATAATTGCATCACTAATTCCTCATATAGCTGATTTGCATCTAACTCTTTGG
```
30   ```
ATAACTCAATGGACCAAAAAAGTTCTGGAATAGCTAACATGATTAAATCAT
TAGAATGATAATGATATACGTTCATCATCATTCCATGAAACTGAAGTGATT
TAATCCTCAAGAGGTCCACCACCAAATTTTTGATATGCTGCTTTTAAGCCA
ATTAAGTCATCGCGATGTGGGACTTTAACATGACCAGGCATAATTTGATAA
GGCTCTCGACCTTTTGAGGCCAAAACAACTGTATCACCTGGTTCAGCAATA
```
35   ```
TCAATCGCGTGTCTAATACCTTCTGCACGGTCATCAAACTCTATATAATTG
TTATGCGTTGCACCTTTAGCTAATTCAGCTGTCAACATTTTAGGATCATCG
TTAGCAGGATTATCTGGAGTAAAAATAACATAATCTGCTCGACATGCTACT
CTCCCCATCTCAGGTGTTTTCGTTAAATCTCTTTCCCCAGCCATACCTACT
AAAAAGATTAGCTTTTGTTTAACAAATGGTTGTACAGCATCAATCAGTTTA
```
40   ```
TTCATACCATCAGCTGTATGTGCATAATCTATAATCAAATCTATAGGAAGT
GATGGATCAAGTACTTCTAAGCGCCCCTCTACAGGTTCAAGCTCAGTAACA
GCATTAATAATTTCATTTAAATTTGTTCCCTTACTCCAAACCGCAATCATA
GCAGCCATAATATTTGAAATGTTAAACCTACCTACATAAGGAGATTTAACT
GGAAAACTTCCAAATGGAGTACAAAATTCAAATTCTACACCTTGCAATGAT
```
45   ```
TCTTTAATGTTTTTAGCCATAAATTGTGCTTCGTGAGTAATACCATATGTA
AATACTTCATAGGGTGTCACACTTGCTAAATAATCGGAGAAG
```

     Sequence 3382
50   step.1002d01.cons.ok
```
CTGTCGTAAAGCTGGTATTATTACTGGTTTGCCAGATGCTTACGGACGTGG
ACGTATTATCGGAGACTATCGTCGTGTTGCTTTATACGGTGTAGATTTCTT
AATGGAACAAAAACTTAAAGACTTTAACACAATGTCTACTGAAATGTCTGA
AGATGTAATTCGTTTACGTGAAGAATTATCAGAGCAATATCGTTCACTTCA
```
55   ```
AGATTTAAAAGAATTAGGTCAAAAATATGGATTTGATATTAGCCGTCCTGC
TACTAACTTCAAAGAAGCTGTGCACTGAGAGATCCCCTCATAATTTCCCCA
AAGCGTAACCATGTGTGAATAAATTTTGAGCTAGTAGGGTTGCAGCCACGA
GTAAGTCTTCCCTTGTTATTGTGTAGCCAGAATGCCGCAAAACTTCCATGC
CTAAGCGAACTGTTGAGAGTACGTTTCGATTTCTGACTGTGTTAGCCTGGA
```

AGTGCTTGTCCCAACCTTGTTTCTGAGCATGAACGCCCGCAAGCCAACATG
TTAGTTGAAGCATCAGGGCGATTAGCAGCATGATATCAAAACGCTCTGAGC
TGCTCGTTCGGCTATGGCGTAGGCCTAGTCCGTAGGCAGGACTTTTCAAGT
CTCGGAAGGTTTCTTCAATCTGCATTCGCTTCGAATAGATATTAACAAGTT
5    GTTTGGGTGTTCGAATTTCAACAGGTAAGTTAGTTGCTAGAATCCATGGCT
CCTTTGCCGACGCTGAGTAGATTTTAGGTGACGGGTGGTGACAATGAGTCC
GTGTCGAGCGCTGATTTTTTCGGCCTTTAGAGCGAGATTTATACAATAGAA
TTTGGCATGAGATTGGATTGCTTTTAGTCAGCCTCTTATAGCCTAAAGTCT
TTGAGTGACTAGATGACATATCATGTAAGTTGCTGATAGGTTTCCAGTTTT
10   CCGCTCCTAGGTCTGCATATTGTACTTTTCCTCTTACTCGACTTAACCAGT
ACCAACCCAGCTTCTCAACGGATTTATACCATGGCACTTTAAAGCCAGCAT
CACTGACAATGAGCGGTGTGGTGTTACTCGGTAGAATGCTCGCAAGGTCGG
CTAGAAATTGGTCATGAGCTTTCTTTGAACATTGCTCTGAAAGCGGGAACG
CTTTCTCATAAAGAGTAACAGAACGACCGTGTAGTGCGACTGAAGCTCGCA
15   ATACCATAAGCCGTTTTTGCTCACGGATATCAGACCAGTCAACAAGTACAA
TGGGCATCGTATTGCCCGAACAGATAAAGCTAGCATGCCAACGGTATACAG
CGAGTCGCTCTTTGTGGAGGTGACGATTACCTAACAATCGGTCGATTCGTT
TGATGTTATGTTTTGTTCTCGCTTTGGTTGGCAGGTTACGGCCAAGTTCGG
TAAGAGTGAGAGTTTTACAGTCAAGTAAGGCGTGGCAAGCCAACGTTAAGC
20   TGTTGAGTCGTTTTAAGTGTAATTCGGGGCAGAATTGGTAAAGAGAGTCGT
GTAAAATATCGAGTTCGCACATTTTGTTGTCTGATTATTGATTTTTGGCGA
AACCATTTGATCATATGACAAGATGTGTATCTACCTTAACTTAATGATTTT
GATAAAAATCATTAGGGGATTCATCAGAGCTGTGCAATGGTTATACTTAGC
ATATTTAGCTGCTATCAAAGAACAAAATGGTGCAGCAATGAGTTTAGGACG
25   TACTTCAACATTCTTAGATATTTATGCTGAACGTGATTTACAAAATGGTGA
CATCACTGAACAAGAAGTTCAAGAAATCATTGACCACTTCATTATGAAATT
GCGTATCGTTAAATTCGCGCGTACGCCTGAATATAATGAATTATTCTCTGG
AGATCCAACTTGGGTAACTGAATCTATCGGTGGTGTAGGTATTGACGGACG
TCCAATGGTAACTAAAAACTCATTCCGTTTCTTACACTCATTAGATAATTT
30   AGGTCCAGCACCAGAACCAAACTTAACAGTGTTATGGTCTACTCGCTTACC
TGAAAACTTCAAAATCTATTGTGCTAAAATGAGTATTAAAACGAGCTCAAT
CCAATATGAAAATGATGACTTAATGCGTGAAAGCTATGGCGATGATTATGG
TATCGCTTGCTGTGTATCTGCCATGAAGATTGGTAAACAAATGCAATTCTT
CGGTGCACGTGCTAACTTAGCTAAAGCATTACTTTACGCTATCAATGGTGG
35   TAAAGATGAAAAATCTGGTAAACAAGTTGGGCCAAGTTATGAAGGTATTAA
ATCAGACGTACTAGATTATGATGAAGTCTTCGAAAGATATGAAAAAATGAT
GGACTGGTTAGCTGGCGTATATATCAACTCATTAAATATCATTCACTATAT
GCATGATAAATATAGCTATGAACGTCTTGAAATGGCTTTACATGATACAGA
AATTATTCGCACAATGGCAACTGGTATTGCCGGATTGTCTGTAGCAGCTGA
40   CTCTTTATCAGCGATTAAATATGCACAAGTTAAACCTATCCGTAACGAAGA
AGGTCTTGTAACTGACTTTAAAATCGAAGGCGACTTCCCTAAATATGGTAA
TAATGACAGTCGTGTTGATGAAATTGCAGTAGATTTAGTTGAACGTTTCAT
GACTAAATTACGTAGCCATAAAACATACCGTAATTCTGAACACACAATGAG
TGTATTAACAATTACTTCAAACGTTGTTTATGGTAAGAAAACTGGTAACAC
45   ACCAGATGGACGTAAAGCTGGCGAACCATTTGCACCTGGCGCAAACCCAAT
GCATGGTCGTGACCAAAAAGGTGCATTATCTTCACTAAGTTCAGTAGCTAA
AATACCTTACGATTGCTGTAAAGATGGTATCTCAAATACATTTAGTATCGT
ACCGAAATCACTAGGTAAAGAAGAAGCAGATCAAAATAAAAAACTTAACTAG
TATGTTAGATGGTTATGCAATGCAACATGGTCATCACCTCAACATTAACGT
50   ATTTAATAGAGAAACATTAATTGATGCAATGGAACATCCAGAAGAGTATCC
ACAATTAACGATTCGTGTATCTGGATACGCTGTAAACTTCATTAAATTAAC
ACGTGAACAACAATTAGATGTTATTTCACGTACATTCCACGAATCTATGTA
ATAAATTTAAGGTGGGAGCAAAATGCTTAAAGGACACTTACACTCCGTTGA
AAGTATGGGCACTGTCGACGGGCCAGGACTTAGATATATATTATTTACTCA
55   AGGTTGTTTGTTAAGATGTTTATATTGTCATAATCCAGACACTTGGAAGAT
TAACGAACCATCAAGAGAAGTGACGGTTGATGAAATGGTAAATGAAATCTT
ACCGTACAAACCTTACTTTGAAGCTTCAGGTGGTGGGGTAACAGTCAGTGG
TGGCGAACCATTACTACAAATGCCTTTCTTGGAGCAATTATTCAAAGAATT
AAAAGCGAATGGTGTTCACACATGCATTGATACTTCTGCGGGATGTGTGAA

```
     TGATACACCAGCATTTAATCGTCATTTTGATGAATTGCAAAAGCATACAGA
     TTTAAACATATTGATAATGATAAGCACATCAAATTAACAGGCAAACCTAAC
     ACACATATTTTAAAGTTTGCACGTAAATTATCTGATATGAAACAACCTGTT
     TGGATTAGACATGTTTTAGTACCTGGTATTTCGGATGATAAAGAAGATTTG
  5  ATAAAACTAGGAGAATTTATTAATTCTTTAGATAACGTTGAAAAGTTTGAA
     ATCTTACCATATCATCAACTCGGTGTGCATAAGTGGAAAAATTTAGGCATC
     CCTTATCAACTCGAAAATGTTGAACCATCTGACGATGAAGCGGTTAAAGAA
     GCTTATCGCTATGTTAACTTTAATGGCAAAATACCCGTAACATTATAGAAA
     GATAAATATAGTAATAGCCTGAGACATCAATCGTTGTCTCAGGCTTCTTTA
 10  CGTTTAATCAGTCGATAAACTGAGTATAAAATATATCTCAATTAACTTTAT
     TTAATAAAAATATGATTCATAAGTTTAAATAGTTTATTTTTCTAGATATGC
     TTTAAATACATCTTTATAAACATCGATGTAACTTAAATACATATCTTTTCT
     AATAAATTCATCAACTTGATGCGCCATAATAGATTCACCAGGTCCAAAGAC
     AGCAAAATCCACGTTATTTTCATTTGTTCCTAGGAAACTAGATGCATCAGT
 15  TGTACCAATCAATGCACTCACAACAATGTCTTCATGTACATAATTCGGTGC
     AATACGTGTAATATTTTGAATAAGAGGATTATCACGATCACTTGCCACTGG
     ATCGTGACTGCTAGGTATATCTACAGTTAAATAATCTTCGCCCACATGACG
     AATGACTTTTTCAAATAAATCCTTCACGAAAGTACTGTCATATTCTGGAAC
     AGTTCTTACATTATATTTAGCTGTCGCTTTATGAGGAACAGAATTAACTTG
 20  TTTACCACCATTGAATACAGAGTTTAACATTACAAATCCAGAGTAGATATG
     TGATTCTTCTTCACCAATTTTTCTGTGGAGATGTTTCTCAATCATTGGAAC
     AGCGTCTAACTCGTGTACTTTATCATGTTCTTTAATATTTTTATATTCTTG
     TTTCATTTCATTTACAAAATCAACTAAAATATCAACAGCATTTGTACCTAG
     GTGTGGCATGGAACTATGTGCAGCCTTACCTTTCGCTGTGACTACACAGCT
 25  CATAGAACCTTTATGAGCATAGTAAGCGATATTACTAGTTGGTTCGCCAAT
     AATTAAACCACTCACATCATCTAAGTAACCTTCATCAGCTAGTAATTGTGC
     ACCATATTGTT


 30  Sequence 3383
     step.1002d02.cons.ok
     GAGATAATCATAAAAATTAGATACTACAGTCATAGTCAGATTAATTGTCTT
     AGGACTAATGCTTTTTTCTTTTCGGTGATAAGAGAGGACATTCTCATATTC
     AAAAGGTGTTTTCATCCATCTTATAAAATCAACAAAGTTATCAAAACTAAC
 35  TTCTTTATAGCATATCTTTTTACTTTCTAAGTAAACAAAAAAATTTTTTAG
     TGCATAGGCATAGGTTTTCTTGGTATTTAAACTTTTCTTAACACTATCCAG
     ATACTTCAAATATCTTACTGCATCTACTATAGGTTCATTATTACCATCTAA
     AATCATAAAATTGGTACCATTCTTAGATTTTACTTCTACTATTTTCATAAT
     AATCTCATTTCTAATTTTAATAGTATAAGTATACACTATTCAAATTCTTTT
 40  TACATCATACATATTAAACTTATTACACGTAATAAAAAAAGACTATTTAAT
     ATTTTTACTTTGTTCATTTTTTCTTAAAAATGATTCATTTTAAACAAAAT
     AGTCTATATTCTTCAACATATATATAACCATTCGCAACTTTACCAGAACAG
     TTTGAAACAATTTATCAATATATGATAGATCAAAATAAAGTGAGCAGACCA
     GTGCTTTCAGACGACCAACTTTCGCAACTAAATATACATTTACATGAAGCA
 45  CTACAACAATCACGCCCAGTCAATATCAAATATTACGAAGAAGGATATATT
     AATTTTATAGAGCTCATTGTACATCGGATTGATTCAATAAATTATGAGATT
     GAAGGTACTGCTCCTCATTCAAGAGAACGTCATAAAGTTTCATTTTTAGAT
     ATTATAGATATTTCTTTTATATAAAGAGTACAAGAGAATTTATTATGAATT
     TCTCATTTTACACTTATATAACCACCAACATACTTAGTGTTATCTAAAGCC
 50  TTTCTGAATGAAATCATTTAATTATTAATTTTAGAATTCAATAATTAATTT
     TGACTGATAATTGGTAAAATCAACTAATAAAAATAAGCCTACACATCTCAA
     TGATGTACAGGCTTATTCCTAAATTATTATAAATTAATTGTTGTTTTCGAT
     ATCAGAAACGATTTCTTTCGTTTTTCTTTCGATATCTTCAAAGTCTTTTTT
     AAAAATATAAGCAAAGGCAACTACACCGATGCCTACTACTACTGTAGTAAT
 55  AGCTAATAAACTAAAGACAAATTTGAATAACCCTTTAATAAAGTTCCACAT
     TAGTTTTCACCTCTATAATTAATATGCATAGCTACATTTCATAATGATAAC
     ATACTTAACTTTATCATTAAATATATACCCTTTTTACAGGTAATTATTCTA
     AATATGAATATAAATTATGATAATATATCAACTAAAGATGCTTAAAATGAA
     ACAATTAAATGATTTCTTTAATCATGGTACACACACAGATTAACTTAGCAA
```

```
        CCAAGATTTTAATTTATCAGTGATTACTTTGTAATTATCAATATTCAGAAT
        GATCACTCCTATTGAGATAATGTCATTATCTTAATTTAGATTGTATGATTT
        ACGCAAATCAAGTTCAATTAATAATAAGTTACTTTTATATTACATTGAAGT
        AAAAATATAAGGAGAGATTATAAAATGCAAAAAGTACGTTCTGATATAATG
5       ACTCATCGTGGTTCTCACTATGATTTAGGAGTAAAGACTGCTCTTTGGTTA
        CAAACCACTCCTTTATTAAAAAATCGAAATAAAGAATGGCGAAAGAGAATT
        CCACGTTTTGATATTGATGTCAAGGAAACCTATGATATATTTCAAATCTAT
        TCGCCACAGATTTGGGAAGAAATAATTGGTATGCAAGATGTATTGAATCTA
        CCTACAAAACAAATGATTTTAAATTTTGGCCATTATCGATTTACTGATTTA
10      AAGGACAGTGGTTGCACAGTATATAAAGGTCGTGATTTTTTAGTCCGAAAT
        TATGATTATCATCCTGCAACATATGATGGTAGATACTTATTATTTCAACCT
        AATGACGGGGGATTATCTCAAATAGGACCGACTTCAAGAGTGACTGGTAGA
        ATGGATGGTATGAACGAGTATGGTTTAGTTATGGCATATAATTTTATGCAT
        CGTAAAAAGCCTGCAAATGGATTTGTATGTTACATGGTTGGTCGGCTAATA
15      CTTGAAAATTGCAAAAATGTAACTGAAGCAATCAAATTTTTAAAGGAAGTA
        CCGCATCGTAGTTCATTCAGTTATATACTAATGGATAGACATTCGAATTAT
        GCCATTGTCGAAGTTACACCTCGATCAATAGATGTAAGGTATGAACATATA
        TGCACAAATCATTTTGAATTGCTTACCCATGAAAATAGAAACTATACAAGA
        GAATCTAAAGAACGCTTAAATCGTGTAATAAATAAAACAACTCCTTCTACA
20      AACAAAGATATCGCATTCAAATTATTTAACGACCCGCAATACGAAATCTAT
        AGCAACCTATTTAAAAGTTGGTCTGGTACAATTCATACTTCACTATATGAA
        CCTAATTCATTAATATCATGGATGGCATTAGGTCAAAACAGTCATCCAACC
        TCAATCAATTTTTCTAATTGGTTAAAAGGAAAGAAATTGAATATAAATTAC
        TTTGAAGGCGAAATAGATACACCATTAACTTTTGCCACATACTAATTAAAA
25      ATTCGACCCTATCTTGATAAATCAAAGGGTCGAATAATTTATATTATTTAT
        TCAATTTTAATATCGCCATCTACAGTGTAAAATTCCAATATATTATCGCTA
        TTTCCCACTTTTCCTTTATCGAAATATCGATTTTTAATTTCTTTATTTCCA
        TGACCAGGATGAAGTTTAAGCAAAGTGTTTTTAGGTTTTTCGCCAAAATGA
        TAATTAATATTACCTCTTTGTGTAGAAGCTTTTATATCATTGTTAGATTTC
30      ATGTCCGTCATAGAAATATCTCCTCTATTAACCAAGAAAACAGTATTAGCA
        CTCTCGTCAATAATATTAACAGCATGTGGACAAGAAAATGTAAAAAAAAAC
        AGAAAGCAATCCAAAAGAAGAACAAAATA


35      Sequence 3384
        step.1002d03.cons.ok
        TCTGCACCATATTTTACTGCTTGGCTCACTTCGTAATCTGCTGCATCCATA
        ATTTTCAAGTCAGCTAATACTTTAGCATTATTAATATTTTCATTTAAATGT
        TGAACTGCAGGTAACCCTTCATTAATTACAATTGGCGTACCAATTTCAACA
40      ATATCTACATATTCTTCAACTTTTTGAGCTAATTTTGCTGCTTCTTCTTTA
        TTTAATAAATCAATGGCTAATTGTAGTTCCAATGAAATTCATCCTTTTCGTT
        TTGGGAAATTCTCTTTTGCTTATAGTTGTATACCCTTCGCTAAACTTATTA
        AACATATAAAATATTTTGTATATAAAACAACTATATAATTATAAACAATCT
        GGCGCTGCTTCATCGTCTACATAAACTTCAACATTAGGGTGTGTGTGTAAA
45      ATGGTCGCAGGTACATCTTCGGTAACCTGTTCATTTAACAGTTTACTTATA
        GCCTCTTTTTTCTTTGGACCAAATGCGAGTAGGATAATCCTTTTTGCTTTT
        AAAATACTTTTTACCCCCATTGAAACTGCTTGTCTAGGAACATCCTTTTCA
        TTGTCAAAAAATCGACTATTTGCTTTTATGGTGCTTTCTGTTAAGTTCACC
        ACATGTGTTTCACTATTGAAGTCAGTCCCTGGTTCATTAAAACCAATGTGA
50      CCATTTTCTCCAATTCCTAAAATTTGAATATCGATTGGCCCTCTTTCATCT
        AATAAATTGTTATATCGTTCCGCTTCTGCTTCTAGATTTTCTGATAAGCCA
        TCAGGAATATGAATGTGATTTTTAACAAAATGAGGATATTGCTCAAATAGC
        ACTTTATTCATGTAGGTATGGTAACTTTGTTGATGACTTGCTTTTAAACCT
        ACATATTCATTTAAATTAAATGTTTCAACCTGCGACACATCAACTTTATTT
55      TTGATTAATAAATTTACTAAGTAATGATATACGTCAGTCATAGTTCCACCA
        GTTGCTAAACCGAGTTTGGCATGAGGGTGCTGCTGTATTTGTTTAAACAAC
        TCGCATGCTACATAAAATGACGCAAGATTTTTTGAATCTAAATTAATAATT
        TTCAATATCAAACACCTCTAATTTCAATATCACAATGATTGATGACTAACG
        CCTTACTCTTTTATATCTCTTAAATTATACTGTAACAAATGATTAAGATTT
```

```
       TAAACATATTAACGCTCTATATGTAATATTCAACGTTATTTCTTGAATTTT
       TCCTAATTAAATCGAGCATTCGTGTACAAATATACATAGCTTTCTAGTAAA
       ATAGATAGATAAACGAGTTTGCTGGTAAATAATATGTTTATCGGTAGTCCG
       TCACTGACATATTGTTTAGTTTAAGGGAGTGAACGTGTTGGAACCAATTAA
   5   AGAACAAGAAGTGCTAGATTTATTAACTTCTTACTCAAATCAGCCTGTTTA
       CCTACACGTTGAAACAACAAATGGTGCTTATGCAAATCATTTCGATCAACG
       CGTATTTAACGCTGGAACATTTTTAAGAAATATTGTCGTGACTTTTGAACA
       TGCACAACTTAAAGGCGGCGACAAAGATCCATATCGTGTAGGTCTTAAATT
       AAAAGATGGTGGCTGGGTTACGTGCAAGGACTTACGCACTATGAAGTTAA
  10   TGAGAATAACGAATTTTTAATTGCAGGTTTTAATTATGAAGGACAATTGGC
       TGCTACAATAGAAATAAGTAAACAGCCATTTACTATATAAAGGAGGTTGCC
       AATCATGACTGATGAAAGACACGTACTTGTGATTTTCCCCCATCCTGATGA
       TGAAACTTTTTCGTCTGCTGGAACTATCGCAAGTTATATTGAAAAAGGTAT
       TCCCGTCACATATGCATGTCTTACCCTAGGACAAATGGGACGTAATCTAGG
  15   TAACCCTCCTTTTGCAACAAGAGAATCTTTACCATTTATACGTGAACGTGA
       GTTAGAAGAAGCATGCAAAGCAATTGGGATTACAGATTTAAGGAAAATGGG
       GTTAAGAGATAAAACTGTTGAATTTGAACCTTACGATCAAATGGATCAAAT
       GATTCAATCACTTATTGACGAAACAAATCCATCATTAATTATTTCGTTCTA
       TCCTAAATTTGCAGTTCACCCTGATCACGAGGCAACTGCAGAAGCTGTAGT
  20   ACGTACAGTTGGACGCATGCATGAATCAGATCGACCCCGTCTTACACTTGT
       AGCGTTTAGCAATGATGCATCAGAAATTCTTGGAGAACCTGATATTCAAAA
       TGACATATCTCAATATAGTGATATAAAACTTAAAGCTTTTGAAGCACATGC
       TTCACAAACAGGACCATTTTTAAAACAACTTGCTAGTCCCGAAATAGATGG
       TCAAGCACAAAGTTTCTTAAAAATAGAGCCATTTTGGACATATCACTTTGA
  25   ATCTTAAATGGAGGTAAAACATGACAGAATTTGACTTATCCACTAGAGAGG
       GTCGTTGGAAACATTTCGGTTCTGTTGACCCTGTCAAAGGTACGAAACCAA
       CTACTAAAAATGAAATGACCGATTTACAAAGTACTCATAAAAATTTCTTAT
       TTGAAATAGAGGAAGTAGGCATTAAAAATTTAACTTATCCAGTTTTAATTG
       ATCAGTACCAAACAGCTGGTTTATTTAGTTTTTCAACGAGTTTAAATAAAA
  30   ATGAAAAAGGCATTAATATGAGTCGCATATTAGAAAGCGTTGAAAAACATT
       ATGATAATGGCATTGAACTTGAATTTAACACACTACATCAATTGTTGCGTA
       CTCTACAAGATAAAATGAATCAAAATGCTGCAGGTGTTGATGTGTCAGGTA
       AATGGTTCTTTGATCGTTATAGTCCTGTGACTAATATTAAAGCTGTAGGCC
       ACGCAGATGTTACTTATGGTTTGGCTATTGAAAATCATACCGTTACACGCA
  35   AAGAATTAACTATTCAAGCCAAAGTAACAACACTATGTCCTTGCTCAAAAG
       AAATTAGTGAATATTCCGCACATAATCAACGTGGTATCGTTACAGTTAAAG
       CATATTTAGATAAAAATAATGATGTCATCGATGATTATAAAGATAAAATTT
       TAGACGCTATGGAAGCCAACGCTAGTTCTATCTTATATCCAATCTTAAAAC
       GTCCTGATGAAAAACGAGTAACAGAACGCGCTTATGAAAATCCACGATTTG
  40   TTGAAGATTTAATTCGACTAATTGCAGCTGACTTAGTTGAATTTGACTGGA
       TTGAAGGGTTCGATATTGAATGTCGTAATGAAGAGTCTATCCACCAGCATG
       ATGCTTTCGCACGTCTGAAATATCGCAAATAAAACATTGAATTTCAACTAC
       AATGCCAGTTGTATTTTTAGTTTATAAGATGCAACTGGCAATTTTTTAGCA
       TCATCAATATCGATGATACATAATGTAACACAAAGAAACTCACCCCATGC
  45   TAAACATTCAATGTTGAATTTATTAAAGGTCGTTACGAATCTAAATAAATC
       AGAAATTCTAGATTTAATTTATTGTAGTGTATTTATGCACTTGTAATTTAC
       ATACATATTATAACGTTTACAACCAACTTTAAGTCATCAGATTAAAGTATT
       TATTAATCACACATACGTGTTTACTAACAAATGAACTTCAAACAAATGCAT
       CAATTCAATTATAGGTTACCTGAGCTATTGACAAACTATTTACATTTAATA
  50   CATTCGACTAGTGAACCAAGAATGTGTCATGCTATGAAGAAAGGTGAAATT
       CATGATTACTG
```

```
       Sequence 3385
  55   step.1002d04.cons.ok
       ATATTGATACTTTCAATCTACTGTTCAACATCTTTTTAGGTAATTGACTCT
       GTATATATGCTTTATGCTTTGTATAAGACGTTGGCATGTATGTTTGGATAA
       ACTTACCTGCATTCCTAAAACGTGGACGAGGAGAGCCGATAGGTTCCTTAT
       ACGTATCGTTAAAATTAATCTCTATTTCCATAACTCACCTCAAAATAATAA
```

```
        TTCGTTAATTGTCATCTGTTGTTGCAGTTCTTCTTTTCTGAACAACTTATG
        TTTACGTTTAAGTTTTTCTAGTTCATCTTTCGTTACTGTTCCTGAGAATGT
        GTTTCTAAAGTGTATGCCTGCATAGTTACCTAGTTTGAATGTATCTTCTCC
        TAACGGCGTTACACTACACATCTCCCAACCGTCAATTTGATACAACATGTA
    5   TTGCTTTTTAAGTCCGTCGATAAGCCCCATCTGGTTGCCTCCATTTCGTTT
        CATTCATGATTAATTCCTGAACTTTTTCATATTCGTCAAATGGTGATATCG
        TTTTGTTTTCTAACAAACGTTTAACTGCCCAGCCTGACTCAATAAGCGTCT
        TAGCTATTAATGGGTCGTTTTGATAATCTTCTCGATACATAACGCCTAACA
        ACTTTTGATATTCAACTACTTTCATGTGAAGAACCTCTGCGTTTCTTGTA
    10  GTACTCAAACTCAACTACACCTGTTCTCCGTCTTTATTCTTTGCGATGTT
        ACATTCAACAATTGACTTGCCTGAGTCATCAACATCATCACGGTTGTAGTA
        ATCATCTCGATATAACAACATAGCTAAACTTGCATCTGCTTCAATTCCACC
        TGCTTCTTTCATGTCAGATAGCATAGGTCTTTTGTCATTTCTTGTTTCTAC
        ACCTCTGCTCAATTGAGATAGCAACACAATAATTGCACCTGTTTCATTTGC
    15  AATAATCTTCAAATCTCGCGATATCTTTTCGATACCATTACGACGATCTAA
        CTTACTGTCTGTCTGCATAAGTTGTAAGTAGTCAATGAAGATAACCTGTTG
        CACATCTTTGTTCTTCATCGCTTGTTTACGTACATCATGTGTAGTAATATT
        GCTTTTATCGTGTATATCTATATCAAGTTTGAGTATTCTGTCTGCTGCAGT
        TGTTAAACGTGTTAATTCATCCGGTTCTAAATCTTTAATTTCTTTGATACG
    20  AGTTAGTTCTATCCCAGTTTCTGCTGATAACATCCTTTTCAATACAGACAC
        GCCAGTTGTCTCTAGACTGAAGAATGAAGTTTTATAGCCTTGAGACGCTAT
        ATTAAGCATCATATTAAGCGCAAACCCGTTTTACCTACTGACGGTCTCGC
        AGCGATTACAATCAACTGTGTAGGTTCTAAACCACCTATTTTGTAATCCAC
        CAGTTTATAACCTGTATTGATTTTTTGTTTTGGTTCTTCGCTATATAATTC
    25  TTCGACAAAGTGATCTACAATTTTTTTAGTCCCACTTTCTTCACTTGCACT
        AATTAAACTGACCTTTTGTAGTTGTTAAGCATTTCATCAAAATTTTGTAT
        ATTCGGATCAGAATTGAATTCTTGTAATACGTTTTGCGTACGCTCTATTTG
        ATAAAGATTGAGCAAATCTTGTTGATATCTTTCAAAGAATCCGTAACCTAT
        AAATTTTGAGTTATACAAATTTGAAATGGTGTCCATATCTAGGAATGACTT
    30  ATCTTTTGTGGTTTTTAAATAGATTTCGTTATGATCTACCTTACCGACTTC
        GAATACATATTCCATAAATGACTTCATACCATCATGTGAGAACATTTCCGG
        TCTCACACGTAACTTCTCAATTATGTCCGGTTTTTGAAGTAAACTAGCAAC
        GATTGTACTTTCGATTTCATGACGTTCATTCATCGTCACTCACTCCAAACT
        CACTTAACTTCTTTCTGAAGTCGTCTAATATCTTTTTACGTTGTGCTACGT
    35  ATTCTGGATCATTTTTCATTCTCCAACGATGTCTAGCAGTTTTTTCGTCGA
        CAGGCTCTTCTTTTACGACTTTGACTTCTTTTCTCATTATGTTTGGAATAC
        TAGGTGGATAAGGATTAGCATCATTGATATATTGCATTACTGTTTTTTTAG
        TCGGTTCATAATCCCCATTTTGGCTCAAAATGTTAACCCATGTTTCTAATT
        TAGGTCTGTCAAAGTCAATGTTGTATACATGTCTAATTGTCTTAATTACTT
    40  CTAGGGCTTGTTGTTTAGTCATAGGCATTAACTTTCATCTCCTAGCTCTTT
        CTCCATTTGAGCAATTACATCATCTGTTACAGTTTTTTTATTTTTAGGTTT
        AATTTTATTTTCAGCCTCTTTTTTATTCTTAACGTTCTCTTTAGCCCAGTT
        GTTTAACACTTTGATTAGGTAACCTGCATGACAACCTTTGTCTTTTGTATA
        ATCAGTAGCTACTTCAACAACTTCATCTGCATGTTGTCCAATATCATCAAT
    45  GGCATATCCTATCTGTTCCATTTGGTTAGGAGTTAAATTATGAGTAAGGTT
        ACTCATGATGTAATTAATTGAGTTTTTGAAGATATCTTTATCTACTTCTTT
        TTCTCTTTCTTCTTCTACTTCTTTATCTTCTTCTATATCTGTTGCGTGACT
        GTCACGTGACTTCACGTGACTATCTAAAAGTTTTTGTTTTTTTCTTTGCTT
        TTGTTTACGCAAACGGTTTTGTTCTCTTATCTTTTCTAAACCTTCGATGTT
    50  CTGATGTTTTTCCCAATTAGATACTTTAAAGACACCATTCACTTCTTCAAT
        CATGCTTAGCTTTTCGAATGTTTGTAACGCTAATCTTATTGAATTGATAGG
        TCTATTAAAATTCGTTAGCTAACATTTCTTCGTTATAGGGTAGACTTTCGGA
        TAACATAATGTATCCTTGTTCGTTATACTTTCCAGCTAATGTCAGCAACTT
        AACCCATAAAGTGATGATTGTATCTCGTTCTGGCAGTGCTTCTATATACTT
    55  GATTTTGCTATCATCGAACATTCC
```

Sequence 3386
step.1002d06.cons.ok

```
        AAACATCACCTTTAAATAACTGATTAATTTGATGAATAATATTAATTGAAC
        TAAAATCATCCGTTAATCTCATGGAATATGCATAAGAATAATTCCAATTTT
        TTAAATGTGCCAAAAAATAATCTAAAGATTCTACAGTTTGCGATATGTCAT
        CAATAATATATCCATTATGCATATGTTTAACATAATCTGTTCTCTTTTTAT
  5     TAATTTGTGGTATACCCGCGCCAATACAACAGATTTGTAAAAAAAGTTTCG
        GCTCTAAAGATAAATCTACTACAACTCTTAATTTTGAAATAACGCTTACAA
        GATCTTCTTCAAACGGTACTGTTTCAATATCAATGATTGTTTCTTTTTTAT
        TCTGTATCACGTCTTCAATTACATCACTTATTTCCTTTTTCTCTTGGTGAT
        ATAAATCATTGAGATGAGCAACTTCATCTATAAGATTTTCCGTAAGATTAT
  10    CTCTACTCTTAGTTAAAATTTTCAACTTATAGCCATCTTTATGTTGAATAT
        ATTGAAATAAGCTGTTTACAATCTCTCGTATTTCAATCTCGTCCAAACCAT
        CTATCCATAATCCAATATAAGTGTCATACAATTGACTACTTTGGTTAGGTA
        ATGACTTCGCATCAAAAGGCGTTACCCTTAAAAGATTAATATCATTCATGT
        TCAATCCTGCGTAGTGTTCAATCATATTTTGATTAGCTTGTGTATCAACGA
  15    GACAATAATATGCTCTAGAAATAGAGTCATATAAATCTTGTGTCACCACTT
        TATTTCTCTCAGTAAAAATAGAAAAACATAAAGACGATGCATCAATAGTGC
        GTGCCATCATTGAATTATGACGTTCATCAGATGCCACTATAACAGAATCAT
        CTTCATTCAATTCTCTTTCTACATATGATTGGAATTTCTCTTCTATCAACT
        CAGCCATAGATGAATACGTAACCCTTTTAAATTTTGATTGGAAATTTTTAT
  20    TAATCGTTACTGTATTAACATTTAAGTCTTCAACAAATATTTCTTCCCCAT
        CTTTTGAAAAATAGTGTTTTTTATTGTTATCACCATCAGGTGTATAAGTAC
        GCACTGCCGATATAAATCCTCTGTCATCGAATACAAATCGTCTTTGAATTG
        TACTATACTTGTAATCCTCTACCCACATCAGGTACCCCTCTTGGCTGAAGT
        GAATTTTAGAAAACGTGTTATCACCTGTAATCGCTTGAATTTGAAAGGGAG
  25    TAAAAATAAATTCAGTGCCTTCTGGCCATGAAAGATCGCGATAATCAATAG
        CTTGAGGCGTTTGATGTCCAACGCCCTGTATCTCATCAAATACAGACCAAT
        AATGACTTTCATACAAATCATATCGATGGAGGAATGTTCTAAGATATGGAC
        TAAAATTTAAAACTATCAATTGATAATCCACATTATTCGAACTATGCATTG
        TCATTAAACTAATCATATCGTCAAAATCTGTATACTGTTTTTTTAGATAGA
  30    ATGGTCTTGAGGTACTTTCCCACCATCTGTTACGGCTATACCAAGCTGGTA
        TAAATCGTTTCATATTCCAAAACCTCCTCCCTACCAATATTTATTTAAAAT
        TTGCTTATAGCTGTCAAAGTATAAGTACGCTTTAATTACTATGGGAGGTTT
        AAAGCACGGAACTGTTCCAATATTTGAAAAAGGAACATCAAGCTATGTTGA
        CCCTATCCCCCAAGCATTGATTTTAACAGCTATCGTTATCGCCTTTGCTAC
  35    AACAGCTTTCTTTTTAGTTCTTGCATTTAGAACATATAAAGAACTAGGCAC
        TGATAACGTTGAGCTAATGAAAGGAGCGCCAGAAGATGATAGAGAGTAACT
        TGCTTGTTTTAACTTTAGTCATACCTATCCTAACTGCAATTTTACTCATAT
        TTATTGGAAAGCGCCCAATAATAAAAAGATATGTTGCACTTGTCGGTACAT
        TGCTAACATTGGTCATTGCATTGATTAACTTAAAGAATGTGCTACGAGATG
  40    GACCAATCAAGGTTGAACTGGGCTCTTGGAAAGCCCCATACAGCATTGTAT
        TTGTAGTCGATATTTTTAGCGCTCTACTCATCATCACAAGTATTATTATTA
        CCATATTAGTCATTTTATACTCATATCGTTCTATCGGTTTAGATAGAGAAC
        GTCATTATTATTATTTCTCTATCATGTTTATGTTGATTGGTATTATTGGCT
        CATTTACAACAGGAGATATTTTCAACTTGTTTGTGTTCTTTGAAGTCTTTT
  45    TAATGTCTTCATATTGTTTACTCGTTATTGGTACTACTAAAATACAATTAC
        AAGAAACAATTAAGTATATTTTAGTCAATGTTGTTTCATCGTCTTTCTTTG
        TCATGGGTGTTGCAGTTTTATATTCAGTTGTAGGAACTTTAAATCTCGCTC
        ATATTAGTGAAAGATTGTCACAACTTTCTGTACATGACAGTGGCTTAGTCA
        ATATTGTTTTTATTTTATTTATCTTTGTCTTTGCCACTAAAGCAGGCGTTT
  50    TTCCTATGTACGTATGGCTACCTGGTGCTTATTATGCCCCTCCAGTAGCGA
        TCATCACGTTCTTTGGTGCACTATTGACTAAAGTGGGTGTATACGCAATTG
        CGAGAACTCTAAGTTTATTCTTTAATAATACAGTAAGCTTTTCTCATTATG
        TCATCCTTTTCTTAGCATTACTTACAATTATTTTTGGATGTATAGGTGCGA
        TAGCTTACTATGATACGAAGAAAATCATCCTTTACAATATTATGATTGCAG
  55    TAGGTGTCATATTAGTTGGTATTGCTATGATGAACGAATCAGGCATGACTG
        GTGCAATATATTACACACTACATGATATGTTAGTTAAAGCTTCATTGTTCT
        TACTCATTGGCGTCATGTACAAAATCACTAAAACGACTGACTTACGTCATT
        TTGGTGGCTTGATAAAAGGGTATCCTATTCTAGGTTGGACATTCTTTATTG
        CAGCGCTAAGCTTAGCGGGTATACCACCTTTTAGTGGTTTCTACGGTAAAT
```

896

TCTATATTGTTCGAGCGACCTTTGAAAAAGGATTTTATCTAAGTGGTATCA
TTGTACTTTTATCAAGTTTAATCGTGTTATATTCAGTCATACGTATTTTCT
TAAAAGGATTTTTCGGTGAAGTTGAAGGATATACTTTATCTAAAAAGGTAA
ATGTTAAATATCTAACAACTATCGCTGTTGCATCTACAGT

5

Sequence 3387
step.1002d07.cons.ok
CAAAATACTTTAATATTTCAGCTGCAACATGATAATCTCTCAAATCCTCAT
CAAAACCAAGAGCAAGATTTGCAGTAACTGTATCATAACCTTTTTCTATCA
ACTCATATGCGCGCAACTTATTAATTAACCCTATTCCTCTACCTTCTTGAG
GTAAATAAATAATCATTCCACCATGTTCGTCAATATATTTCATTGACGCTT
CAAGTTGTGCCCCGCAATCACATCTTTGACTATGAAAAATATCCCCAGTCA
GACAAGCAGAATGCATACGTACATTAGGATTGCTTTTTAAATCTCCTTTAA
CAATAGCTACGATTTCTTCATCGCTATAATCCGTTGTAAATCCATACATAT
CAAAATGACCAAAATCAGTTGGCATCTTGACCTTTGCCTTAAGATTAACAT
TAAGTTCAACAGCCTTACGAAAAGCAACCAAACTTTTTATAGTAATCATTT
TTAAATGGTGGCGTTCTTTAAATGACTGGAGATCTTCACCCTTAGCCATTG
TCCCATCATCATTCATAATTTCGCAGATTACTCCAGCTGGTTGTGCTCCTG
TTAACCGTGCCAAATCTACGGCAGCTTCAGTATGACCATTACGTGTTAACA
CACCATTCTCTTTTGCTATAAGTGGAAATAAGTGCCCCGGACGATGAAAAT
CTTCAGGATTAGTATTTTCATCTATGAGTGCTCTAGCCGTTTGTGTACGTT
CATGTGCACTGATTCCTGTAGTAGTTTTATAATGATCAATGCTTACAGTAA
AATGTGTGCCATAAATATCAGTGTTATTTTGCTCCATAGATTGTAGTTTTA
ATCTTTCAGCTATAGATTTATCAATTGGTGCACAAATCAGACCACGACCCT
CTTTAGCCATAAAATTAATGGTATTATCATCCATCCATTCCGTAACAGCTA
CAAGATCTCCTTCATTTTCTCTATCTTCATCGTCAACTACAATAATGCTCT
CTCCATTTCTTAAAGCCTCTATAGCCAACTCAATTGTATCGAATTGCATGT
GATATTCCTCCTAAAAACCAAATGCTCTAAGCTTTTCTTCTGTTAACTGCG
ATTGATTTTGATTCATGATGTTTTCAACATATTTGAATAGTACGTCAGACT
CCAAGTGCACTTTGTCGCCCACTTTTTTAGATGAAAGAATAGTAGATCGAC
GTGTTTCTGGTATAAGATGTATATCAAAAGTATAATCATGTAAATCAAATA
CAGTAAGACTAACTCCGTCTACAGTTATAGAACCTTGCTTTACCATTTGAT
TCAAAATATTTTTAGTTGTTTTAATAGAAATAATTTTTGAGTTAGCAGTTT
CATTGATTTTAGAAATTGTTCCAAGCTCATCAACATGACCTAACACGAAAT
GTCCACCAAATCTCCCACTTCCACTCATGGCTCTTTCGAGATTAACTTCTG
TATTACGTTGAACACTTCCAAGATATGTTTTGTTTTCAGTCCCTTTGATGA
CTTGAACTGAAAAACTTGAGTCAGTGAAATCTATCACAGTTAAACATGCAC
CGTTAACACTTATTGAATCACCAATATGCATATCAACTAAAATGTTTTGTG
CTTTAATTTCAAGCGTTCTTACTGATTGTTCAGAGCGAACTTGTTGTACAG
TACCTATTTCTTCAATGATACCTGTAAACATAGACATTCACTTCTTTCGTA
ATTTCAATTTTAAATTTTGATTAATTAACTTGGAATCAACAATTTCAAATT
GAGTTGCTTCAGGCAAATCAATGACCTCGTCAGTCTTATAAAATTGATGTT
TGCCAGAACCACCAATTAATTTCGGGGCTATATATAAAATGAGTTCATTTA
GATGTTTGGATTGGAGAAATTGAGATGTAATATTTGGGCCTGCCTCGACTA
GCAGTTTCCCAATACCTCTTTGATATAAGTCTTGTAATATTGTCGTTGTAT
CACAATTACTAATATTTATTATTTTAATAAAACTTTTATTTGTTTTTAATT
TTTCATTTTCAGTGTAAATCCATATCTCCGATGCAGTATCTTTAAATATTT
GTTGATTAAAATCGAGTTGACCTTTCTTAGAAAGAATAACTCGAATCGGAT
GCTTTCCATCAGGAACCCTGGTTGTATACAATGGATTGTCTGCTTCAATGG
TTCTACGCCCAGTAATAACTGCATCATGCTCATGTCTTAATTGATAAACAT
CTTCTTTAACTTCTTTGTTTGTTATCCACTTACTTTCATTAAAGTCTGTTG
CTTGTTTACCATCTAGACTAGATGAGACCTTTACAGTTACTTCTGGAACTT
CGTTTCTTTTAGCAGTAAAAAAGTCACGGTATAATGCAGCTGCATTTTCAT
TATATTGAAATTCAACCTCTATACCAGCTTCTCTCAGAATCTCGTCACCCT
TACTTACTAAAGTAGTATCTTTAACAGCATAGATGACCTTAGATATGCCCG
CTTCAATGATTTTATGCACACAAGGTGGTGTTGAACCATGGTGTGTGCAAG
GTTCTAATGAAACGTATATGGTAGCACCTTGGGTATTTAAACCTGCCATTT
CAATAGCTTGTACTTCGGCATGTTTATCTCCCTTTTTTAAATGTGCACCTA

```
AACCTACAATCCTACCGTTTTTAACAACAACGGATCCTACTGGTGGATTAA
CACCTGTTTGTCCATTTACCATTTTTGCTAGTTGAATAGCATCATCCATAA
ATCTACTCAATTGATCACCTCAAAAAAAACCTTATACCTATATTGTAGGCA
TAAGGAGATATCTATATCTTTAAATAAGCCTATTAGCATACTTAAAATTGC
ACATAAGTATACAATACGCTTATCATAACTACACGAATTTGTGTAAGTCGA
TAATTCAACTTCTTTCTCCCATCCAGACTCTAACTGTCGGCTCTAGATTCA
CACTAGATCAGCCACTAACATCAAAAACAATATTAGTGGGTCGCAGGCTCA
ATTTACTGCCGGTTGGGAATTTCACCCTGCCCCGAAAGAAATATATGAAAT
TGTTATGATTAATTGAGTCCAGTAGTTCTTCACTACATTGTAAATCTTACT
ATATTTAAATTTAAAATACAATTTTATAGCTTACTTTATTAAAAGATTTTA
TGAGAATATATTAAACTCACAAACACGCTTAATGCATATTACTTATTTCTA
TTTTCCTGTGTTATATATAATTGAGATACTGTATCAACAACTTGATTAACA
ATCATCTTTACGCCATTTCTAAGTTTATTGACACCGTTGGTCATTTGTCCA
ATTGCAATCACATTTTTTAAGGTTCCATATCTGGGGCTTATGACTTGATTT
GTTTCTGGGATAATTTGAATGCCACCCATAGGATGACGTTGAACAATTTGT
CTATTTTCTAAGTTTAAAATTAATTGATCATCCTCATCTAATTGAGAAAGA
TGTGTTTTAGAGCCCGTTGCATTAATAACGATGTCAAACAACTCATAATTT
TGCGTAGTGTCGTTATACTTCAACATTAAGAGGAATCCTGCACAAAAATGT
CTCCATTCGGTAGCCATTTTTGTTAAAAACATTGATTTGACAGTGTCAAAT
AGAATTCCTTTTTATGCGCTGAAATTTGTTATATGGAGACTGAATGTA
```

Sequence 3388
step.1002d08.cons.ok

```
GATTTAGAACAGAATGTTACAGTTACACAAGTTAATGATGCATTTAAAAAT
GCCGATTTATCAGGTGTTCTTGATGTTGAAGAAGCTCCTTTAGTTTCTGTA
GACTTTAACACAAATCCTCATTCAGCAATTATTGATTCTCAATCTACGATG
GTTATGGGACAAAATAAGGTGAAAGTTATCGCTTGGTATGATAATGAATGG
GGTTATTCGAATAGAGTTGTTGAAGTAGCTGACAAAATTGGACAATTAATT
GATGATAAAGCAATGGTAAAAGCCATTTAATTAAATATTTATACTTACACT
AAACCTGAGATAGCCATATTATCTCAGGTTTATTTATATACAAATAGCCCT
AAACAAAGTTGAAGCACCATTGATGTTATTATAGATAACCAATTAGCTTTA
GACATATTTCTAGATACTTTCTTTATAAAAAAATGATCCATAGTCATTATT
TATTAATCAAGTTGAGATGATTTTGAAATGAGGTGATTTTAACGAATCATC
TTTCAAATTATAATTTAATAATGGTACTTTTGGTCGTATTTATAAAATTGT
GTAGTACATTTAAATGGTTGAAATTCATTTAATATGGTTAATCACGTTAAA
ATGTATTATTTTTAATAAATAACATAGCATTCTAAAACCTTTATTTATAAG
ACAATTTTTCCTTTTTTACACATAGTGGTATTAACAATGTATCATTACATT
CTTTCGTTTATAGTTGCAATTGACCTTTCTTTTTTACACAGTACTTTTATA
GTATAATGTACGGTGATTAAACTTGGGAGGGTGACATATATGAAATGCCCA
AAATGTAATTCTACACATTCCAGAGTGGTTGATTCAAGACATGCAGATGAG
GCCAATGCGATTAGACGTAGAAGAGAATGTGAAAATTGCGGAACGCGTTTT
ACAACATTTGAACATATTGAAGTTAGTCCATTAATAGTAGTGAAGAAAGAT
GGGACTAGAGAACAATTTTTAAGAGAAAAAATATTAAATGGTTTAGTAAGA
TCTTGCGAGAAACGACCAGTACGTTATCAACAACTTGAAGACATAACTAAT
AAAGTGGAGTGGCAACTTAGAGATGAGGGACAAACTGAAATTTCTTCTAGA
GAAATTGGAGAGCATGTTATGAATTTGTTAATGCATGTTGACCAAGTTTCC
TATGTAAGATTTGCATCTGTATATAAAGAATTCAAAGATGTTGATCAACTC
TTAGAGTCAATGCAAGGTATCTTGAGTGATAATAAACGGAGTGATAAATAG
ATGGGGTTACAAACCTATGAATATGGTCTAAAACCACAAGATGGATTTGAG
GTGATTACACATTTCGAATTCACCTCACAACATTTAGATATTTTAAATCGA
CTATTCACCCCTTTAATCGGAGTTGAATCAATTGGACTCTATCATTTTATG
AGTCAATTCATAGATAAAAGTCAACAACTCGGGTTAACGCATTATATATTC
ATGAATGAACTAAAAATTAACTTATTAGATTTCAGGGAGCAAATGGACAAT
TTAGAGGCTATTGGATTGATTAAAAACATTTGTAAGGCATGAAGAAAAGTAC
TCTCACTTTGTTTATGAGTTAATTCAGCCTCCAACAGCCTATCAATTTTTT
AATGATCCTATGTTATCAGTATTTTTATTTAGTGAGGTTGATAAAAAACGT
TATCAAGCACTTAAATCTTATTTCGAAAAAGATGAGAAAGATTTAAGCAAA
TATCAACAGACAACTAGAAAATTTACAGAAGTATTCAACGTACCTAAAAAG
```

```
    GTCAATGTTTCTGATCAAATTAATTTAAAGCAAATCAAACACTATGATGGT
    ATAGATTTATCTAATGAAACTTTTGATTTTGAAATGTTGAGACAGATGTTG
    AACCATCATTTTATTAGTAATGAAATTATCGATAAAGAAGCTAAGAATTTG
    ATTATACAACTTGCGACACTTTATGGAATTACTGAAGATGGTATGAAAAAT
5   GTTATATTAAGTTCCATTACCAGTGCACAACAATTATCTTTTGAAGAAATG
    CGTAAGAAAGCTAGAACTTATTACCTGATTGAACATGATAATCAATTACCA
    AAATTAGAGCATCAAACAAATAAAATTAACGATGAAAAAAAAGATCGACAA
    GCGGAAGATACAACAAATGATTGGTTACAACTGCTAGATGAAACAAGTCCG
    ATTGATATGTTAGCAAGTTGGTCTGATTCGGAACCTACACAGTCGCAAAAA
10  AGTATGATAGAAGAATTGATTAACCGTGAAAAAATGAATTTTGGTGTAATC
    AATATACTTTTACAGTTTGTTATGTTAAAAGAAGATATGAAGTTGCCAAAA
    TCTTATATTTTTGAAATTGCTTCCAACTGGAAGAAAATTGGTATTTCAAAT
    GCCAAACAAGCATATGAATATGCATTACAAGTTAATCAACCTAAAAATTAC
    GAAACACATTCTAATGATAAACGACAGAACAATCGTGGAAGACAAAATCAA
15  TTTTTATCCAAAGAAAAGACACCTAAATGGCTTCAAAATAGGGACGATCAA
    GAAGAAATAAAGAAATAAATGATGACACTCTCGAAGAAGATCGACAAGCA
    TTTCTTGAAAAGTTAAATCAAAAGTGGAAGGAGGAAGATAACTAATGAAAA
    GTTTCAGAAATATCATGGGCGATTCTCAAAATCTAGATAAACGTATACAAA
    AAATAAAACAAAATGTAATCAATGATACTGACGTTAAACATTTTCTTGAGA
20  AAAATCGTAGTAATATAACTAATGAGATGATAGACGAAGATTTAAATGTTC
    TTCAAGAGTATAAAGATCAACAAAAAGTTTATGATGGACATCGCTATGATG
    ATTGTCCGAATTTTGTAAAAGGACATGTTCCTGAACTATATATTGAAAATG
    AAAGAATCAAAATTAGATATCTACCTTGCCCGTGTAAAATTAAACATGATG
    AGGAACGATTTGATTCACAACTTATTACATCTCACCATATGCAAAGAGATA
25  CACTTCATGCAAAGCTCAAAGATATTTATATGAATAATCGAGAGAGACTTG
    ATGTAGCAATGGCAGCTGATCAAATCTGTACAGCAATTACTAACGATGAAA
    AAGTAAAGGGGTTATATTTATATGGTCCTTTTGGTACAGGAAAATCATTCA
    TATTGGGTGCTATTGCAAATCAACTTAAATCGCAAAAGATTTCATCAACAA
    TTGTATATTTACCAGAATTTATTCGCACTTTAAAAGGTGGCTTTAAAGACG
30  GTAGTTTTGAGAAAAAATTACAACGTGTGCGAGAAGCTAATATTTTGATGT
    TAGATGATATTGGCGCAGAAGAAGTCACACCGTGGGTAAGAGATGAAGTGA
    TTGGTCCTTTATTACA


35  Sequence 3389
    step.1002d09.cons.ok
    CAATGGCTTTGGGAAACTTCACTTTATCAAAAATGGTTGATTTCTTTTATT
    TTAATTTGCAATATATTTGGTACGATATACGGTTATCATTGGTATAGTAGT
    CAATTAGCTAACACACCAAATTATTTTAAATTATTTGTTCCAGATAGTCCC
40  ACTGCTACATTATTTTTATGTTTATCATTATTATTCATACTATTTAATAAG
    CGTTATGCCCTGATTGACGCTTTAGCTTTTATAACTCTATTTAAATATGGT
    AGTTGGGCTGTAATTATGAATATTTTGATGTTTATAAAAATGAATGATATT
    GCCATAAACGGACTTATGTTGCTCATTTCTCACGCAATTATGATATTAGAA
    GCTATCTATTTTTATCCTCGTTTTAAAATATCTAAATTGGCTGGATTGATG
45  AGTTTCATATGGGTGACGATCAATGCGTAATAGATTACATATACGGACAA
    TATCCCTACTATGATTTTATCGCCAAACATTTAATTGAAGTAGGGGTATTG
    GCTTATAGTCTCACTATCATTTCGTATATTTTATTTTTAAAATTACAAAAG
    TGGTTGAAAGTTAAAACATTTGATTAAAGACAAATTTGACAGTAAAATGAA
    AATAAAGGAGAGTGAGTTTAATGTCTTTAATATCAATGATTATATATTTCG
50  TAATATTAATGGTCATTCCAATGTGGGCTCAACACAAAGTTAAATCTAATT
    ATGAAAAATACTCACAAGTGAGATCCACAAGCGGTAAAACTGGTCGTGAAG
    TTGCTGAAGAAATTTTACATGCTAATGGCATTTACGATGTTGACGTCGTTA
    AAGGCGAAGGTTTCTTAACTGATCACTATGACCCAAATAAAAAAGTTGTTT
    GTTTATCTCCTGCAAACTACGATAGACCATCAGTTGCTGGTACTGCAATTG
55  CAGCCCATGAAGTAGGACATGCAATTCAACATCAGCAAGGATACGCACCAT
    TAAGATTTAGAACAGCATTAGTACCATTAGCTAATATTGGTAGTTCTTTAA
    GCTACATTATTATTATGGTAGGTATTGTTTTAACAGCTCTCGGTAGTGTCT
    TTGGCTCTACAGTTCTATGGGTTGGTGCTGGTTTAATGTCACTTGCAGTTT
    TATTCTCTATCGTGACATTACCTGTTGAGTTTAACGCTAGTTCAAGAGCGA
```

TGAAACAAATAACAGCTTTAAATATCGTTAATGAGAAAGAATACAAACACG
CGAAAAAAGTTTTATCTGCAGCAGCAATGACATATGTGGCATCTACTGCAG
TTGCGTTAGCTGAACTTGCTCGTATTATATTAATTGCAAGATCAAGTGATT
AATCATATATAACAACGCAGTTCCTGGTAATCATTACAATGGATTTGTAAT
5   GGTTACCTTTTTTTATTTGAATACGTGACAAAATCCTATCTTTTAATCCAT
ATTAAACATTGCACTTGAAATGCAAGTCACTTTATTTTGTAATATACTATA
TAAGAATAGCGATAAAGGAGTACTTCTATATGAAATCTATGGAGCAAATGC
AACAAGAAGTTGATGACTATATCAGTCAGTTTAAAGCTGGTTATTTTTCTC
CACTTGCTAATCTAGCAAGAATGACAGAAGAAGTTGGAGAATTAGCACGCG
10 AAATTAACCATCATTTTGGAGAGAAAAAAAGAAAGATACAGAAGAAGATA
ATACAATTAAAGCTGAATTGGGAGACAATTTATTTGTTTTACTGTGTATTG
CAAACTCTCTAAACATAGATATGACAGAAAGCTTTAATGACACGATGAATA
AGTTTAATACAAGAGATAAAAATCGCTTTGAAAGAAAATAATTTAAGAAAG
GATGGGCACAAATGAAAATCGGTATAACGTGTTACCCGTCCATGGGTGGTT
15 CTGGCATTATTGCTACAGAACTCGGTATTAAAATGGCAGAACGTGGCCATG
AAGTTCACTTTATTACCTCAAACATACCCTTTAGAATACGCAAACCTTTAC
CTAACATGACGTTCCACCAAGTTGAAGTCAATCAATATGCCGTATTCCAAT
ATCCACCATACGATATTACATTAAGTACAAAAATTTCTGACGTTATACAAG
AATATGATTTAGACATATTGCATATGCATTATGCTGTACCTCACGCTGTAT
20 GTGGTATATTAGCGAAACAAATGTCAGGTAAAAACGTCAAAATTATGACAA
CACTACATGGCACTGATATAACTGTGTTAGGTTATGACCATACTTTACAAA
ACGCGATAAAATTTGGCATAGAACAAAGTGATATTGTAACAAGTGTTAGCC
ATTCTCTAGCACAGCAAACTTATGAATTATCAATACTAAAAAGGAAATCA
TCCCTATATATAATTTTATAAGGGAAAATGAATTCCCAACTCGGCATAATG
25 AAGAATTAAAAGATTGTTATGGTATTTCACCTGAAGAAAAGGTATTGATAC
ATGTTTCTAATTTCAGAAAAGTAAAACGTATTGATACAGTGATTGAGACAT
TTGCAAAAGTTCATGAGAGTATACCATCCAAGTTGATACTTTTAGGAGATG
GTCCAGAATTAATCGATATGCGACATAAAGCACGAGAACTTGATGTTGAAA
CACACGTACTCTTTTTAGGCAAACAAAATGACGTAAGCGCATTCTACCAAC
30 TATCTGATTTAGTACTACTCTTAAGTGAGAAAGAAAGTTTTGGATTAACTC
TCTTAGAAGCAATGAAAACAGGCGTCTTACCTATAGGGAGTCATGCAGGTG
GTATTAAAGAGGTCATCAGACATGAAGAAACTGGATTTATAGTAGATATAG
GGGATAGTACACAAGCTGCAAAATATGCTATTAAACTTTTATCAAATCCAG
AGTTATATCAAAAAATGCAATCACAAATGCTGAAAGATATTGAAGCAAGAT
35 TTAGTTCAGATTTAATTACTGACCAATATGAAAACTATTATCGAAAGATGC
TAGAACAAGGTGAGAACAACAATGAGTCATGAGATATTTGAACAAGCAAAA
CCTATTTTAAAAAAGATTCAAAATAAAGGTTTTAAGGCTTATTT

40 Sequence 3390
    step.1002d10.cons.ok
    TTAAAGATTTGAAAGAACTTTCTGACGCAGATATCAAAAGTTTTGAAGAAC
    GTTTAGATAAAGTCGATAATCAATCAAGTATTGACCGTATTATAAATGATG
    CAAAAGATAAAAATAATCATTTAAAATCGACAGACTCTAGTGCCACATCAT
45 CAAAAACTGAAGATGACGATACATCTGAAAAAGATAATGATGATATGACTA
    AAGACTTAGATAAAATACTGTCGGATTTAGATTCAATTGCTAAAAATGTTG
    ATAACCGTCAACAAGGTGAAAATAGTGCTTCTAAACCTAGTGACTCAACAA
    CCGATGAAAAAGATGATTCAAATAATAAAGTACACGATACAAATGCTAGTA
    CACGCAATGCAACTACTGATGATTCTGAAGAGTCGGTTATTGATAAATTAG
50 ATAAAATCCAACAAGATTTTAAGTCTGACTCTAATAATAAGCTTTCTGAAC
    AAAGCGATCAGCAAGCATCACCATCTAATAAAAACGAAAATAACAAAGAAG
    AATCTAGTACGACAACAAATCAATCCGATAGTGATAGTAAAGACGATAAAA
    GTAATGATGGTCGTCGCTCAACATTAGAACGTATAGCATCAGATACTGATC
    AAATTAGGGATTCAAAAGATCAACATGTCACAGATGAAAAACAAGATATAC
55 AAGCAATTACACGTTCACTACAAGGTAGTGATAAGATTGAAAAAGCACTTG
    CTAAGGTACAATCTGACAATCAATCACTAGATTCTAATTATATAAATAATA
    AATTAATGAATTTAAGATCACTAGATACAAAAGTAGAGGATAATAACACTT
    TATCTGATGATAAGAAACAAGCGCTTAAACAAGAAATTGATAAGACTAAGC
    AAAGTATTGACCGACAAAGAAATATTATTATAGATCAACTCAATGGTGCTA

GTAATAAAAAACAAGCAACCGAAGATATCTTAAATAGTGTTTTCAGCAAAA
ATGAAGTAGAAGACATAATGAAACGTATTAAAACAAATGGCCGAAGTAATG
AAGATATTGCTAATCAAATTGCCAAGCAAATTGATGGTCTTGCATTAACTT
CTAGTGATGATATTTTAAAATCAATGTTAGATCAATCTAAAGATAAAGAAA
5   GTTTAATTAAACAATTGTTGACGACACGACTTGGTAATGATGAAGCAGATC
GTATTGCTAAAAAATTGTTAAGCCAAAACTTGTCGAATTCTCAAATCGTAG
AACAATTAAAACGTCATTTCAATAGTCAAGGAACAGCTACAGCTGATGATA
TATTGAATGGTGTGATTAATGATGCTAAAGACAAAAGACAAGCGATTGAAA
CAATATTACAAACCCGTATCAATAAAGACAAAGCTAAAATTATCGCTGATG
10   TTATTGCGCGTGTACAAAAGGACAAATCAGATATCATGGATCTCATTCACT
CTGCGATTGAAGGCAAGGCAAATGATTTATTAGATATAGAAAAAACGAGCAA
AACAAGCTAAGAAAGATTTAGAATATATTTTAGATCCTATAAAGAATAGAC
CATCCTTGTTAGATCGTATTAACAAAGGTGTCGGTGATTCTAATTCAATAT
TTGATAGACCAAGTTTACTTGATAAACTTCACTCAAGAGGATCTATTCTTG
15   ATAAATTAGATCATTCGGCACCGGAGAATGGATTATCTTTAGATAATAAAG
GTGGCCTTTTAAGTGATCTATTTGACGACGATGGTAATATCTCATTACCAG
CGACAGGTGAAGTCATCAAACAACATTGGATACCAGTGGCTGTTGTACTCA
TGTCATTAGGTGGAGCGCTCATCTTTATGGCGCGTAGAAAAAAACACCAAA
ATTAATTTAAAAATTATATATAGTAATAGGCATCATCAATCTTATGTGTAG
20   TTAAACTATTTCACAGTTAGTGATTGAGGTGGTGCAACGCTTACTCAAAGT
TTATAACAAGAGTCAGTATAAAAGAGTAGGACTAAGAATTTAAAATCAATT
CTGTCCTACTCTTTTTTGACATGCAATAATGATTGTTATATATAATGAATA
ATATACTTTTAGAAATCACTTGAGTAAATTATAAAAATGGACTTTTCAATG
TAAAGTTACTCTAATATAGTTATTGAATATAAATAAGCACAAAGATGTATT
25   ATAAATTTGAATACATCTTTGTGCTATTATTTTGTGGATGGTACCTCAAAT
AATTTTAAATTTTTACTATTTACCTTTAATGTAATCTAAATCTTGAGCATC
TTCTGCATCAATAGTTCTGAATGAAATGACACCTAATCCTTTAACATTTGA
TTCAGAAATAACAATTGAACCGTCGTCATTAACTTTTTCAACGAAAGCAAC
ATGACCATACTGTGTATCAGCACCTAATTGTCCAGCTTCAAACACAACTGC
30   AGTATGATTTTTAGGTGTGTGCGTTACCGTATAGCCTTCACTTTCAGCACG
ATTATTCCAATTATGAGCATCACCTAAGTCACCAGAAATGGATGCATCAAA
TTGATTCATACGGTGGTACACATACCAAGTACATTGGCCATGTGGGTAAGG
TGATGTACCGGAAGTTTCAGTAAATGGTTTGAAGTCATTACCAGAGACATC
CGTACCTATTGATTTGTTGTATTTCTTCATGTTAGGCATTTTTTCTTTGTC
35   AAAAGATGTTAAATGATAATGTTTAATAATACTATTTAATTTTTTAGAATA
ATTAGGATCTGTGGCGTATGAGCGTGACAGATGTGAAGTAGCATCTTTATA
TGATAGAGCTTCACTCTTCCAAGTTGGTTTATAAATTGACGGATTACCATC
GATACCATGTTTGATTAAATCTGCATAATCTTCAAGAGATTGTTTAGTACT
TGGGTATTTACGGAAACCTGCTTGGATACTGAACATATGATTACTGCTATC
40   AGCTTCTAAAGTATTAAAAGTTACAGATTGTCCTTTGTAGTCACCTTTGAT
TCCAAACAAGTTATGATTTGGTGATTGTGCAAGTGAACTTTTTCCAGAGTC
AGATTCTAAAATAGCTTGAGCAATCATAACTGATGCATATATATCTTGGTC
TTTTCCAATCTGATGCGCATCTTTAGCAATTGATTGAATGAAATCACGTGT
GTCTTGACTGTCAACAACATTAAATGAATCAGAAGATAAGTCACCATTGTC
45   TAATTCAGGTAGTTGTTGGAAAAGACTTGTTGAACGTGTATTTGAGCGTTT
AATGTCATCCTCAAATGATTGTGCAGGTTTTGATTGATGTTTTAATTCATC
ATCTGTTGGTAATTGTGGATTCTTGTCGGCATTATCACTTTGAGATGTTTT
AGTGTGACTTGTATTCTTGCTCTTATTGTAATCTTTTTCTGTTTTCTTAGC
GTCCTGACTATATTCATCTAATATAGAGTCTAATGCTGAATCTGTAATACT
50   AGAATATTTCTTGTTACGTTCTGAGCTATCATCATTCGCTTGAGTTTGTTG
TTGCTCAGAAGAAGATTGATTTTTGTTTGTATCT

Sequence 3391
55   step.1002d11.cons.ok
TCATAGGGCTTAGATAATGCTTTATCCAAACCCACCTTTCCATCATTCTCT
ATATCAACAGTATAATTTTCATGAGTTAACTCAAGCTCTATAAATCTAGCA
AGATTTTGCTCATCTTCTACAATTAAAATATTTGTCATATTTGCACCTCAC
GCTACATCTTAAACAATTAACTCGTATTTCTAAAGATGTAATTCAAAATTG

```
TAAATGATTTGTTATGAAACTAGTATCATTTTACCATGAACCACGATAATA
GATAGTCGAATAAGTAGTCTGTATTTAGAAAAAGTCACTCTAGTATTATCA
TTGTATTTAAGTTAAATGGAAAAAAATATATATTTCTAATTGACAATCATT
ATCAATCATGTATAATGATAATTGAAAATGATTATCAATACCAATTGAAAA
ACATTCCCCCCACATACAAGTTGTTCTTTTGGATTGGTCATTTTCAACTAT
CCCCTTTTATATGCCCGTAAAAGACTAACGTTAAGAAATGACGTTTCAATA
AAAGCAGTAGACCTTTGACACTTGAGGTCTGCTGTTTTTTCAATGATAATA
TAATGAGAGAAAGTAAAATAATAAGTGACTAATAATCTTATCAAAATTTTT
AGAACGCTCATCACTTTTGATTAATGAAATAATATATACAAAAGAGTCTGT
AACACATGAAGTGCACAGACTCTTTCCTACGATTAGTCAAGTACTTATAAT
TAAAATAAGCATTACTTAACTAAACCGTTCATTTTAATTTTTTTGAATAAG
GTGTTAAACTATATTTGATTGTACGTTTGGCTAATGTGGTGACATTTGATA
ATCGTACAATTTATGATTCATTACAAGATAATGAATTGTTTTTAATAATCG
ATTTATACAAGCAATGATGGCAGTCTTATGAGGTTTCTCATTAGGCTGCTT
TCTTAGTTTGTAGTAATAATCGACGACATGATTGTCATAATGATGCTGCCC
TCTTATTATATTCATAATCACCCAAAATAAAAGTTTTCTCGCTTTTTTATT
ACCACGCTTGTTGATGGTATCTCTACAGTGTGTATGACCTGATTGATATCG
TTTGATATCAATGCCAACAAAAGCATTGAGTTGTTTATTTGATTTAAATCG
CTTAATATCACCAATCTCCCCAATAATCATAGCTGTGCTTAGCTTACCAAT
ACCAGGTATCGAATGAATATTTTCAAAATAATCGAGTTGTTGTGCTAATTG
AATCATGGCATCATCTAATTGTTTGAGATGATGAATAGATTGTTTTAATTG
TTGAATAAGTAAGCGTAATTTTTCGACTAGAAAGGAATGTCTATCGACATT
AGGATAGCTTTCTTGAGCAATCACCCTTAATTGAAGTGCATATTTTGTAGC
TTTATCCATTGACATTCCCTTATCTGTAGAATTGAATATATGTGTAATCAG
TACCTCCTTGTCGATATCAAGAACCATGTCTGGATGAGTAAAGATTTCTGC
GATGTTGAGTGCAATGATTGAATATCGACTACTAAATAATCTTTCTAAACC
AGGGAATGTTTGATGGAGTAATTCAAGGATCTGAAATTTAAGTCGATTTTG
TTCATTCTCGATTTCTAGATGAAAACGGACGCGTTCTCTTAATTCAAAGAA
TATTAACTCATGTATAGGTAAGTTGTCTGTTTGTTTAAGCGTCGGTCCTAA
ACAAGCAAGCTTATGAGCATCTGCCTGATCAGTTTTCCATGATCTTAGAGC
GCTCGTTTTAAATTTGGCTTCTAACGGATTCATTTGAATATAGTTAATTTG
ATTTACACAACAAATCGTTCCATACTTCTTGAATAGATACCTGTAGATTC
AAAAATGAGTTGTGGGTGGTCTAAGTCATTCAAATACTTGAGTAAATAGTT
GTAACCATTTTTATTATTTTGGATGAAAAACTCTTTTTGGAATTTTTCATT
TTTATAATGTGCAACTACACTACTTCTTTTACTAATATCAACACCTAAGTA
ATCGATAAAAAAACCTCCTTTGAATAATTGAGAAGCTAAAAACTTTACTTA
ACCTTTCTCATTTCATTTTCCTATACACGGTTTCAAGAACCCAACATACTA
CAAACGAATTTCAAAAGGCGAGAGTAAAGCTGACTTGTTTTTTATACGGAT
TTAAAATCCAAGAGTCTGGACAGTCTACTTCTCTCTATAACTATAAAAAAT
AGCTATGAAAAAATCTATCGTCATAGATTTCTTCATAGCTAATCTTAGTAT
GTTTTTATTTTCATAATTAATTGTTATGTTGCTAATTTTGGATGGTTTTCA
ATCTAATCAGCTTCGTTAGTCACATAATATTGAATGAACTAGGTTCAAACT
AAAAGTAGTTGAGTGTTATCAAAAGAAAAAGAAAGACCCTACTTATAATG
TAAATATCAAAGGGACTGAAAAAGCTAAAGATGAGAAATTTAAGTTTATTA
CGATAAAAGACTATTTAAACTTGCTAGATTATTTCAAGAAAAGAGATGAAG
AAAGTTATGTTTTGCTATATCTATTAGGCATTACTGGCGCAAGATATAGCG
ATGTCATTTATATGACTTACAAAGATCTAAACAAAGCGAATGGCATAATTC
ATTTGCCTGGAACGAAAACAAAGAATTCAAAACGTGATGTAGAAGTTAATT
CAAAAGATATCATGCACATAAATTCAAAATTAGCTAAAATGCCGCGTAGAA
TTGATGGCAAGTTATTCTCGGTTAGTCATACATCAGTAAGTAAAGCGTTTA
GAAAAGCAAAAGAAGTGATAGGATTAAACGATAATAATATAACTCCCTATT
CACTCAGACATACGCACACATCTTACTTACTATCTAAAGGCATACCAATCG
AGTATATAAGTAAACGTTTAGGTCACGCTACTATATCACAAACGTTAGACA
CGTATTCACATTTATTAGAAGAACATAAAAAAGAGCAAGGCCAACGTGTCA
GAGAAATATTCTCTTGACACTTATTTGACACTTACTCTCTCGAAAGCCCGT
CATATCAACGGTATAGTACGGAGAGTTTCATTACTTCTTAATGGTGGATTC
ATTTTGTAGTTTGAATCATCGCCAGGAACATCATTTTCAGTTAATAATAAA
TGATGTGGTGTAACTTCTGCTGTTACATGGATACCAGCTTTTTTAGCGTCT
CTGAT
```

Sequence 3392
step.1002d12.cons.ok

5   TCAAGATTGGAAGTTACGTCATGTTCTAATTGAAATGCATGGTAATAATGG
TAGTATCGATAACGATCCTCCAGCTGCTATGCGTTACACAGAAGCTAAACT
TAGTCAATTATCAGAAGAACTATTAAGGGATATTAATAAGGAAACAGTATC
ATTTATTCCAAACTATGATGACACAACTTTGGAACCAATGGTATTACCAGC
GAGATTTCCTAATTTATTAATTAATGGATCTACGGGGATTTCTTCAGGATA

10 TGCTACTGATATCCCGCCGCATAACCTCGCCGAAGTAATACAAGGCACATT
GAAGTATATCGATCAACCTGATATTACAATTAATCAACTGATGAAATATAT
CAAAGGGCCTGACTTTCCTACAGGTGGTATCATTCAAGGAATAGAAGGTAT
AAAAAAAGCGTATGAGACCGGTAAAGGAAAGGTTGTCGTGCGTTCACGAGT
AGATGAAGAGCCTTTAAGAAGTGGACGTAAACAATTAATTGTGACTGAAAT

15 TCCGTATGAAGTGAATAAAAGTAGTTTAGTTAAAAGAATTGACGAATTACG
TGCCGATAAAAAGGTTGATGGTATTGTAGAAGTTCGAGATGAGACTGATAG
AACTGGATTACGAATTGCAATCGAATTAAAAAAAGATGCTAATAGCGAATC
AATCAAAAACTATTTATATAAGAATTCGGATTTACAAATTTCATATAATTT
TAATATGGTTGCTATTAGTGAAGGTCGCCCTAAGTTGATGGGATTACGTGA

20 AATTATAGAAAGTTATTTAAATCATCAAATTGAAGTGGTTACAAATAGAAC
GCGTTATGACTTAGAGCAAGCTGAAAAACGTATGCATATTGTGGAAGGATT
AATGAAAGCTTTATCTATACTTGATGAAGTTATTGCATTGATACGTAATTC
TAAAAATAAAAAAGATGCTAAAGATAATTTAGTTGCAGAGTATGACTTTAC
TGAAGCTCAAGCAGAAGCTATTGTCATGTTACAGCTGTATAGATTAACAAA

25 TACTGACATTGAAGCTTTGAAAAAAGAACATGAAGAGTTAGAAGCTTTAAT
AAAAGAATTAAGAAATATCTTAGATAATCATGAGGCACTTTTAGCAGTAAT
TAAAGATGAACTAAATGAAATTAAAAAGAAATTTAAAGTGGATCGACTATC
TACAATCGAAGCTGAAATTTCCGAAATCAAAATTGATAAAGAAGTTATGGT
GCCTAGTGAAGAAGTGATTTTAAGTTTGACGCAACATGGCTATATAAAACG

30 TACATCTACACGTAGTTTTAACGCAAGTGGTGTGACTGAAATCGGTTTGAA
GGACGGCGACCGTTTATTAAAACATGAAAGCGTGAATACTCAAGATACTGT
TCTTGTATTTACAAATAAAGGTAGATATTTGTTTATACCTGTTCATAAATT
AGCCGATATCCGTTGGAAAGAGCTTGGTCAACACATATCACAAATTGTGCC
AATAGATGAAGATGAAGAAGTGGTAAATGTATACAACGAAAAAGATTTTAA

35 AAATGAAGCCTTTTATATTATGGCTACAAAAAACGGCATGATTAAGAAAAG
TAGTGCTTCACAATTTAAAACTACTCGGTTTAATAAACCACTCATAAATAT
GAAGGTTAAAGACAAAGATGAACTTATTAATGTCGTTCGATTAGAGTCTGA
TCAGTTAATTACTGTTCTAACCCATAAAGGCATGTCATTAACTTATTCAAC
TAATGAATTATCGGATACAGGCTTAAGAGCAGCTGGTGTTAAATCAATTAA

40 TCTTAAAGATGAAGACTATGTTGTTATGACAGAAGATGTGAACGACTCAGA
TTCCATAATAATGGTTACACAACGTGGTGCTATGAAGCGTATTGATTTTAA
TGTTCTTCAAGAAGCTAAACGCGCACAACGTGGAATTACTTTACTAAAAGA
ATTAAAGAAAAAACCGCATCGAATTGTGGCAGGTGCAGTAGTTAAAGAAAA
TCACACGAAATATATTGTATTCTCTCAACATCATGAAGAATATGGTAATAT

45 CGATGATGTACACTTATCTGAACAATATACTAATGGATCATTTATTATTGA
TACTGATGATTTTGGAGAAGTAGAAAGTATGATTCTAGAGTAAAAGTATAT
GCAATCACAAAATAAAATGATAAAATAAAAATTAATAATAGCAACTAAAAT
GAGCACAAGGGATTTATGTCTAATTTGAACTATCAAAGTGATAATATACAT
TTATTGTTAATATGAATGTTTAATCTTTGTTAATGAAATTAAAGCATAAAT

50 CTCTTGGTGTATTTAAAACGATAAATTTAAGTGGATGAGAGGGATTCGTTT
TGCAAGATTTTGATAACTTAATTCCTGGCTGGTTTAAAAACATTTGTTCAAG
TCGGGAATGACTTAATTTGGTCTCAATATCTTATTGGATTATTATTAACAG
CAGGTTTTTTCTTTACAATTAGTTCTAAATTTATTCAACTCAGAATGTTAC
CAGAGATGTTTAGAGCATTAACTGAAAAGCCAGAAACTTTAAGTAGTGGTG

55 AGAAGGGTATTTCACCATTTCAAGCTTTTGCGATTAGTGCTGGGTCAAGAG
TAGGAACTGGAAATATTGCCGGTGTTGCAACTGCTATTGTTCTTGGTGGCC
CCGGTGCAGTCTTCTGGATGTGGATTATTGCTTTTATTGGTGCAGCTAGTG
CATTTATGGAAGCAACGCTTGCTCAAGTTTATAAGGTACATGACAAAGAAG
GTGGATTCCGTGGCGGACCAGCCTATTACATAACAAAAGGGCTAAACCAAA

AATGGCTTGGAATTGTATTTGCTGTTTTAATTACAGTTACATTTGCTTTTG
TATTTAATACTGTTCAAGCGAATACAATTGCTGAATCATTAAATACACAAT
ACAATATTAGCCCGGTAATTACTGGAATAGTACTTGCAGTTATTACAGGTA
TTATCATCTTTGGTGGTGTTCGTAGCATAGCTACACTATCTTCACTTATTG
5   TGCCTATTATGGCTATTGTTTATATAGGTATGGTTTTAATCATTTTATTAC
TCAATATAGATCAAATTGTACCTATGATTGGCACTATTATTAAAAGTGCAT
TCGGAGTTCAGCAGGTTACTGGTGGTGCTGTAGGAGCTGCTATTCTTCAAG
GTATTAAACGTGGTTTATTCTCAAACGAAGCTGGTATGGGATCTGCACCTA
ATGCTGCTGCTACATCTGCTGTGCCCCATCCCGTTAAACAAGGTTTAATTC
10   AATCATTAGGTGTATTCTTTGACACTATGCTTGTTTGTACAGCTACAGCAA
TTATGATTTTATTATATTCTGGTTTGCAATTTGGTGATAGCGCGCCTCAAG
GTGTAGCAGTTACGCAATCAGCGTTGAACGAACATTTAGGTTCAGCAGGAG
GTATTTTCTTAACTGTAGCAGTTACCTTATTTGCATTTTCATCTGTTGTAG
GTAACTATTACTATGGACAATCCAATATTGAATTTTTATCTAACAATAAGA
15   TGATATTATTTATTTTTAGATGTTTTGTAGTACTTTTAGTATTTGTAGGTG
CTGTTGCTAAAACAGAAACAGTTTGGAGTACTGCCGATTTATTTATGGGTC
TTATGGCAATAGTAAATATCATATCAATTATAGGTTTGTCGAATATTGCGT
TTGCAGTGATGAAAGATTATCAAAGACAG

20

Sequence 3393
step.1002e02.cons.ok
TGGAGCTGGTAATTTATATATGACGCCACATGATATGGGCAAATTAATTTA
TACGTTACAACAAAATAAAATCTTTAATGCACGTCAAACTCGACCTATTTT
25   ACATGAATTTGGAACTCAAGAATATCCAGAAGAATATAGATATGGTTTTTA
CATAACTCCGTATTTAAATAGAGTCAACGGGGTATTCTTTGGTCAAATTTT
TACTGTTTACTTTAATGATCGGTATATTGTCATTTTAGGGACGAATGTAAG
TAATACACCTGGATTAGTGAGTAATGAAGACAAAATGAGACACATTTTCTA
TAATATTCTTGACCAGAAAAAGCCTTATAATACAGCAGGTGTTAAAGTTGA
30   GTAATTGTAATTTTGTTAATAATGGATTAGACTCGTTATTCCATATTTACA
AGGTTAGTCTTCAATAAATTATATATAACGAATGCCCGGAACTTCAATCTG
AAGTGTAACCGGGCATTTGTTATATGATTAATAAAATAACATTCAAATTGC
AATTTAATTATTTACAATAAGAAAGGGCGTTTAAGAATAAATCCTCATACT
GTCTTTCAACAGGACTTAAAACAACTGTACAATTGGGATGCTCCATATTAA
35   AGTCTACCACTGTGGCTCCCCTTGTAAAGTTGCCATTTAATTCTATTTGAG
TATACGCCTCCTTGACATTAAAGGCTTCTGGATCCAACAAATATAATATAG
TGAATACATCGTATAATTTAAATCCTATTTCAAAATCTTCACTCTTATAGT
GCTGAAATAAATTATATAACATGTTTGAAGTTGCATTTGTATCTTTGAAGT
CTTTTACAAAATGGTGAGTAAACAATGCTTCACGAGCCAAATCAAGACCAA
40   TCATAGTCAAAGGTAATCCAGAGTTAAATACAATTTGAGCAGCTTCTGGAT
CACAATATATATTAAATTCAGCTAAAGGCGTTACATTACCTCTACCGGTAC
TACCACCCATTAAAACAATTTCCTTAATAAATGGTTGAACTTCGGGATAAC
TAGTTAAAAGAATAGCGATATTTGTTAGTGGACCTATGGCAATCAAGGTTA
AAGGTTCTTGAGTATTTACTAATAGATTTCTCATTGCTTCAACTGCATGAA
45   TTGATGTTAAATCATCTTGATTTATTTGTGGAAACTCGTAACCATCCATAC
CAGACTCACCATGAATTGATGTAGCTTCAAAGATGTCATTAATCAATGGTT
GGGATGCCCCTCTATGTACAGGAACAGAACTATTAAAAAACCTTTTTAGCT
TTAATGCATTTGCTGTCGTTTTCTCAATACCTACATTACCATTCACAGTTG
ATATCATTTTAACGTCAAATTGAGGGTGTGAAAGTGCGATACTAATTGCTG
50   TAGCATCATCTATACCAGGATCAGTATCAATTATAATGGGAATAGACATAT
CATCATCTCCGTATTTCATAATAATTAAAGCATAACATGGTATGGGATTGT
TACGATAACGAAATAAAAAGAGCCTGAGACTTTTAAGTCTCAGGCTCTTTA
CATATAAGGTGGAAATGAATCAATAAAGAACAAATATATTCAAATTCATTC
TTTATTGAAATCAACCCTGGAAAACTCATCTTTATAAGTGAGAAGAGTGTC
55   CACCTTGCATTACCCAGTATGTTCCGATAACAGTTACTAAAGTAATGATAA
TTGCAAAGATAACTTTGAACGATTGTAAACGTCCATCTTTACCTTCAGTTA
AATGCATGAACATTAATAATTGAAGGGCAGCTTGAATGAAAGCAAAACCAA
AGATGATAGTTACCTTAGCATGGAATGTCATATTAGTGTATAGAGTTACAA
AAACTGCTAAAAGCGTTAGTACGATAGAAGCAATAAAACCTACTGTATGTT

904

```
     TTACGATTGTATTCATCCGCTATACACCATCCCTATCATATATACGGCAGT
     AAAGATGAAGACCCAAACAACATCTAAGAAGTGCCAGTATAAACTTACTAT
     AAATAATTTAGGAGCATTGTATGAATCCAAACCACGAGTGCCGATTTGAAT
     TAACAAACAAATAACCCAAACAATACCTAATGATACGTGTGCACCGTGCGT
 5   ACCTAGTAGTATAAAGAAACTAGACCAGAAGGAGCCAATAGTTGGGTTAAC
     ACCTTCAGAAGCATAGTGTGCGAATTCGTAAATTTCGAAACCTACGAATAC
     AAGACCTAGGATAACTGTGATAATCATCCAAAACATCATTAAGTTTTGTTT
     TTCTTGTCGCATGTAATAAATTGCAATACCACAAGTATAAGAACTAATTAA
     TAATGCAAAAGTCATTATTAAAATCAAATGTAATTCGAATAAGTCGGTAGT
10   TAATTTACCGCCATATCCGCCACCATGTTGTAACGTTAATAACGTTGCAAA
     TAGGGTACCGAATAACGCAAATTCAGCTGTAAGGAAAATCCAAAAGCCAAG
     TTTATTTAATTCGCCTTCGTGTGTACGAGAATCAATAGTATTTGCATCATG
     ACTCATGACTTACAGCCTCCCTTTCTTTAATTCGAGCTTCTCTTAAACGAG
     CTTCAGTTTCTGCAACTTCTGAAGCAGGGATGTGGTAACCATGATCAATTT
15   GGAAACTTCTCCAAATCATAGTAATGAAGATACCTGCTAAACAGATAAGTG
     CTGGAACAATAGATTCGAAGATTAAGAAGAAACCACCAATAGTCATAAATA
     TACCCATCCAGAATCCTACTGGAGTATTGTTTGGCATATGAATATCTTTGT
     AGTTATGGTTGTCTAAATAATGACGACCATGTTCTTTCATATCAACGAATG
     TATCGTAGTCATTCCAATCAGGAGTGATAGCAAAGTTGTATTTAGGTGGAA
20   TAGCTGATGCTGTAGACCATTCTAAAGTACGACCAAGTCCATCCCAGTTAT
     CTCCAGTAGCTTCACGTGGAGCTTTGATATGACTATAAACGATACTTGCAA
     CTAGGAATAAGAATCCAATTGCCATCAATACTGCACCGATAGTTGAGATGA
     AGTTTAGTAACCACCAACCATCAGAAGGCATGTAAGTGTATAGACGACGTG
     GCATACCATCTAAACCTAGAATGAATTGTGGTAAGAAACAAACGTTAAATC
25   CGATCATGAAGAACCAGAAGCACCATTTGTTTAATGTTTCATTTAACTTGT
     AGCCCATCATTTTTGGATACCAGAAGATTAAACCAGCTAAGCAGGCAAATA
     CTACACCAGTAACCAATGTATAGTGGAAGTGAGCTACTAAGAAATAAGTGT
     TGTGATATTGATAGTCAGCTGATGCCATTGCAAGCATTACACCAGTAACCC
     CTCCTAATAAGAAGTTAGGGATGAATGCTAATGAGAATAGCATAGGTGACT
30   CAAATGTAATTCTACCTTTGTATAATGTGAGCAACCAGTTAAATAGTTTAA
     CTCCGGTTGGAACACCGATTAACATTGTTGAGATAGAGAAGAATGAGTTAA
     TTAACGCACCATTACCCATAGTGAAGAAATGGTGAACCCAAACTAAGAAAC
     TTAAGAATGCGATACCTGCAGTTGCCCAAATCATACTTTGATGACCGAATA
     AACGTTTACGGGCAAAAGTAGGGATGATTTCTGAGTACATACCGAATGCTG
35   GCAAAATAACGATATAAA


     Sequence 3394
     step.1002e03.cons.ok
40   GCATTAGCACTTAAAATGGATAACAAGCCAAATATTGCTACCGCAACAGTT
     GGAGAAGGCAGTTCAAATCAAGGTGACTTTCACGAAGGTATGAACTTTGCT
     GCAGTTCACAAATTACCTTTCGTCTGTGTAATAATTAACAATAAATATGCG
     ATATCTGTACCAGATTCACTACAATATGCTGCTGAAAAGTTATCAGATCGT
     GCATTAGGTTACGGTATGCATGGAATACAGGTAGATGGAAATGACCCAATT
45   GCAGTATACAAAGCGATGAAAGAAGCAAGAGAACGAGCGCTAGCAGGTGAA
     GGTCCAACATTGATAGAAGCTGTCACTTCACGTATGACACCACATTCATCT
     GATGATGATGATACATATCGTACAAAAGAAGAAAGAGACCTATTGAAACAA
     GAGGATTGTAATATAAAATTTAAAACGGCCTTACTCGATCAAGGCATCATA
     AACGAAAATTGGTTGAGTCAATTGGAAAAAGAGCATAAAGAACTCATTAAT
50   GAAGCTACTAAATCTGCTGAAGCAGCACCATATCCTTCAGAAGAAGAAGCT
     TTGACATATGTTTATGAAGAGGGAGGTCAACGAAATGACTAAATTATCATA
     CTTAGAAGCTATACAAAATGCACAAGACTTAGCACTAAATCATTTTAGTAA
     TGCATTTATACTTGGCGAAGATGTAGGAAAAAAGGGTGGCGTTTTCGGCAC
     AACCAAAGGATTACAAAGTAAGTATGGTGATGAACGTGTAATTGATACACC
55   TCTTGCGGAATCGAATATTATTGGTACTGCTATTGGTGCAGCAATGCTAGG
     TAAGAGACCTATTGCCGAAATACAATTTGCAGATTTCATTTTGCCTGCTAC
     AAATCAAATTATAAGTGAAGCAGCTAAAATGAGATACCGTTCAAATAATGA
     CTGGAATTGCCCACTGACTATCAGAGCACCTTTTGGTGGTGGAGTTCATGG
     TGGATTATATCATTCACAAAGTGTTGAAAGCATTTTTGCTTCAACTCCTGG
```

ATTAACTATTGTTATTCCTTCATCACCTTATGATGCTAAAGGTCTTCTATT
GTCCTCTATAGAGTCTAACGATCCAGTCTTATACTTTGAACATAAAAAAGC
ATATCGTTTTCTAAAAGAGGAAGTTCCCGAAACATATTATACTGTACCTCT
AGGTAAAGCAGATGTTAAAAGGCCAGGCGAGGACATCACTGTATTTTGTTA
5   CGGATTGATGGTGAATTACTGTTTACAAGCTGCAGATATTTTGGCAAATGA
CGGCATCGATGTTGAAGTAGTCGACTTAAGAACAGTTTATCCACTAGATAA
AGCAACTATCATTGAACGCTCTCAACGTACTGGTAAAGTTCTTCTTGTTAC
TGAAGATAATCTAGAGGGAAGCATTATGTCTGAAGTATCTGCAATTATAGC
TGAAAACTGTCTGTTCGATTTAGATGCGCCAATCATGCGATTAGCTGCACC
10   GGATGTCCCATCTATGCCATTTTCACCAACATTAGAAAATGAAATTATGAT
GAACCCAGAAAAGATACAGGACAAAATGCGTGAACTCGCACAATTTTAAGG
AGGGTGTCAATTGGATATAAAAATGCCTAAGCTTGGTGAAAGTGTGCATGA
AGGTACGATTGAACAATGGTTAGTATCAGTAGGAGATCATGTAGATGAGTA
TGAACCATTATGTGAAGTTATTACAGATAAAGTAACAGCTGAAGTGCCTTC
15   AACAATTTCTGGAACAATAACAGAATTAGTGGTTGAAGAAGGACAAACTGT
CAATATTAACACGGTGATTTGTAAAATCGATTCGGAAAATGGTCAAAATCA
AACAGAATCGGCAAATGAGTTTAAGGAAGAACAAAATCAGCATTCTCAATC
AAATATAAACGTGTCACAATTCGAAAATAATCCTAAAACTCATGAAAGTGA
GGTGCATACAGCCTCTAGTCGCGCAAATAACAATGGACGATTTTCACCAGT
20   TGTCTTTAAATTAGCTTCTGAACATGATATTGATTTAACACAAGTCAAAGG
AACTGGTTTTGAAGGTCGTGTTACTAAGAAAGATATTCAAAATATTATTAA
CAATCCAAACGATCAAGAAAAAGAGAAAGAATTTAAACAAACAGATAAAAA
AGATCATTCAACGAACCATTGTGACTTTTTACATCAATCCTCAACTAAAAA
CGAACACTCACCATTATCAAATGAACGTGTCGTACCAGTTAAAGGTATTAG
25   AAAAGCTATCGCACAAAATATGGTTACTAGTGTCAGCGAAATACCACACGG
TTGGATGATGGTTGAAGCTGATGCAACGAATTTGGTTCAGACTAGAAACTA
TCATAAAGCTCAATTTAAACAGAATGAGGGTTACAATTTAACTTTCTTTGC
GTTTTTTGTAAAAGCTGTTGCAGAGGCTTTAAAAGTAAATCCATTACTCAA
TAGTACATGGCAAGGAGATGAAATTGTTATCCACAAAGATATTAATATCTC
30   TATTGCTGTTGCAGACGATGATAAGTTGTATGTGCCAGTCATTAAAAAATGC
AGATGAAAAATCAATTAAAGGTATCGCGCGTGAAATCAATGATTTAGCTAC
TAAAGCAAGATTAGGAAAATTAGCACAAAGTGATATGCAAAACGGTACATT
TACGGTTAATAATACTGGTTCTTTTGGTTCTGTTTCTTCAATGGGAATCAT
TAATCATCCACAAGCTGCCATTTTACAAGTAGAATCAGTCGTTAAGAAACC
35   TGTAGTTATAGATGATATGATTGCAATTAGAAATATGGTTAATTTGTGTAT
TTCAATCGATCATCGTATTCTCGATGGTGTTCAAACGGGAAAATTTATGAA
TCTTGTTAAGAAAAAAATAGAACAATATTCTATTGAAAACACTTCTATTTA
TTAATTACAAACAATGATTACCAATTAAACTTCATATTAAAGTTGAACGCT
TTAAGGAGGATAATTTCTTCTTAAAGCGTTCTTCAGTATTCGTTTTAAATT
40   TCAACTAGAATTTTTCAATCATCATAAATTAATCCTGTTGAACAATGTGTT
GTTAATTAGTAGAATAAAATAAGTAAAATTAACTTATTATGGAAGGTGACA
AATCATGGATTTAAACTTTGATTTATATATGAATGATGTTGTTGAACAAGC
ACGAAACGAAATTGAGCATGCTGGTTATCATCAATTAACTTCTGCTGAGGA
TGTTGATCAAGTTTTACAACAAAAGGGGACATCTTTAGTCATGGTAAATTC
45   CGTA

Sequence 3395
step.1002e04.cons.ok
50   AGACCGTCTACAATATACAGTTTTTACTGTTGAAACACACACTACTTTTTT
ACATGAATTAACACTTGGTCAAGAATTCAATATAGAGCTATACCTTTATAA
TTACGATGACAAACGAACTCATTTCTTTCTGAGAATGTTAATTGATAATCA
AGAAGTTGTTGCAACCAATGAGGTCATGATGTTAGGAATAGATAGGACACA
ACGTCGTGCTGCACCATTTCCAAAACATTATCTGAATGCTATACAAGATTA
55   TGCTCATAAGCAAGAAAAAATTGAATGGCCACCACAATTAGGCCACTCAAT
TGGAATACCATATAAAGGAGAATAATAATGTCTATAAAAGAAAAATTAGTA
ATAGAGCTCATGAATTTAAAAGACACTGTAAAATACGAATTGTCAGAATTA
AATGAACAACAATATTTGTTTATGAACGGCCCAGCACCTCAGTTGATGAAG
CGTGCCATGAAAATGAGCTATATTCTAGGCCAAAAAGAAGCGATTGATCAC

```
         TTTCTATTACTCATCGACACTTGCCATGAGAATGTCTTACTAGATAACTGC
         CAAGACTATCAAAAACAAATTCAACATAAACAATTCTCTTTACAAGAAGAT
         ATTATTCACCAAATAAATCAGTCAAAGGAATTCGAATCATTTTTATCAACT
         TATTATGTAAATAAAGGTAAATATGATATCACCAGCAAAATAAAAACATTA
      5  CTTGAATAAATACTATTGTTTAAATAAAATGATTATCTCATACATCACTAG
         CTGATAAGCATAGCATTGGTAATTGACTTCATCCTTCCTTTTAGAACAACA
         TTCACATCACTTTTAAAATACAGTTAAAAATCTATTCAATTCTTACTCTAA
         GATTGATTTAGTGATGTGTTTTAACACCTTGTGAGATTGGCGTATGGGATA
         GATAGGAAATCATGCACCATTTTAGGATACTCTTTAAACTCAACATCGTT
     10  CCCACATTCCTCTAGAGCATGAGTAAATGCATGCATGTCTGGGCTTAATAT
         TTCTCTGCCACCTCCATACATATATATAGGTGGCAACCCTTTTAGAGTACC
         GTATAATGGTGATATACGCGCGTCTGATAAAGGTAAATCATTCGTCCAAGA
         TTTCATGAGTTGATGAACACCAAATCGACTAACTAAGATATCATTTTCTTC
         CAACGTTTTCGTAATATTAGGATTAGTCAAAGTGGCATCTAAAAGCGGTGA
     15  AAGTAAAAATAACTTTCTAGGCACTTCTTGATTATCATTGATAAGTGATTG
         AACAAAACTTAATGCTAGTCCTCCACCAGAACCATCCCCCATCATAACAAT
         GTTGTTCGCACCAACTTCTTCAACTAATTGATTATAAACGTCACGAATCGC
         TTTGAACGTTTCTAAGTAGTGATAATCTGGTGACTTTGGATAGATAGGTAA
         AACGACCTCATGCAATGTATTTAAAGTTAGTTTATCTAAAAGTCTCCAATG
     20  AAATGGAGAAGGTTGCAAAGTATTATATCCACCATGCAGATAGAGTATCTT
         TTGATTTTTCTCATGTCTAAAATTAAATCTAAACACTTGCATATCATTAAG
         AGTAAGTTTATCTAAATTTGACTTTACATTAAGTGTCGCAGGTTGCTGATG
         TTTTTTACTATTTTCGATACTTCTCTTTTCTAAAAAATGTTTAACTTCATC
         ATCATTACTAAAAAATATCGATCTATTGTGAAGTATATATTTATTAACCAC
     25  GCGATTCATCATTCTATTTTTCATTAATCAATCAACCTACCTCTATTTAAA
         ATTTTAAATTTAAGAAATAATTTGTTTAAATATCTATATACCTTTGCTAAA
         AAAAGATATAATGATAAGTATGTCATTTTAAAGAAATATTTTTTAATCAAT
         ACTTAGGATACTAGATTAAACCCAATTATAACAGACATAAAATTAAAAAGA
         TAAACTTAGTAAATTTGCATCTAAAGGAGGCATTTAAATGTCTGAAACATT
     30  ACAAAACCAAAGACAATTCAACTCTCAATTTGAACATCAAGAGATCCATCG
         TGGAAAAAGATATGGTAAAAAGAAACGCTCATGGGTAAGTCTCATTATTCA
         AGTTTTCGTTTTAGTATTAACTGCTATTACTGGCTATAGTATGCTTAAACA
         ACCTATATTTAAAATTTCATTTGTAAATGAAACTATAAATTTCCATCAATT
         AAGAAATTTTCAAGATACAGTTACGCAAATCGGTAATTTGAATTTAGGTAA
     35  TATTGATCAATTACAACAGTCTGTCGATAATCTCATCATTATATTTAATAT
         TTTCTTTGTCCTATGTCTGATTAGTTTGTTTATAACCGTAATTACAATTAT
         TTTTAATCGCACAGCACTCAAAGTAGTCAATATATTACTTTTTAGCTATTA
         TGCTAGTCATTACATTGTATTTCAGTTATATCATACATACAATTGCTCAGA
         AAATTTCCGAATCTCTAAAACAATATTATTTAACAGTATCACCAGAACAAG
     40  TATTAACTGAAGCAGATGCGATTCACAATGCACTGATACTCATTGGATGTA
         GTATTGCGTTATTGATTGTAAGTTTGTTTTTCCGTAATCGTTTACCACGTA
         TTAAATAAAGATGATTTTAATGTTCTTTTAA
```

```
     45  Sequence 3396
         step.1002e05.cons.ok
         CCATAGCAACATTTTCAGTTATGAATTGATTACTTTGTATGGGATGTGTAA
         CTTCTATCACTTTCATAATATTATTTCTCCTTTAATTAAAAACTTTTTTAG
         GTTTTATTTCATAAGGTTTATCTGGATGAACTATATATATTGCTAAAACTT
     50  TTTGAGTGCTACTATCGACAAATATAAAAGGTTCTTTAACATTTTGACTTA
         ACACTTTTTTATGAAATTTTTGACCGTTACAGATTTTCTTTTTAAAATTTG
         AATCTTTCACTTGGAATGATTTCAGACCTTTTAAGCCATATTCTATAGGAA
         ACAATTCATTATGTAATGAATCATGCTCATGTAATTCTTTAATTTGATCAA
         TCGTTAAACTACTTTCTAATTGAAAACCGCCAGAAGCAATTCTAGTTAGAC
     55  GTGACATATGAGCTGGAAAACCAAGTTTAAGTCCAATATCTGTAGCTAAAG
         TTCTAATATAAGTTCCTTTACCACATGTTACTTCAACTTCAAAATGACATG
         TCTGCTCCTGAAAAGTAACTTCAGATATTCTATGTATATCTTTAATAAAAA
         CTTGTCGTTTAGGGCGTTCGACAGTTTCATTATTTCTCGCATATTCATATA
         ATTTTCTTCCATTTACTTTAACAGAAGAATACATAGGCGGAATTTGTTGAA
```

```
     TATGCCCCTCAAATTGCTGCAACACTTGGTCAATCGTATCTTCATTAATAT
     CATTCTTATCAACAGCCCTAGTTTCTAAAATATCTCCAGTTTGGTCTTCAG
     TAGTCGTACTCTTTCCTAGCGTTATCATAGCGTGATAAGTTTTTCCCATTT
     CCATGATATAATCACTGACTTTTGTCGCATCGCCTAAACAAATTGGTAACA
  5  CACCATTAACTTCAGGATCTAATGTTCCTGTATGACCAATTTTTTTCATTT
     TTAAAATTTTACGTAATTTAAAAACGACGTCGTGACTTGTTAAACCTCGTT
     TCTTAAATACCGGTAGTATGCCATTATACATGTTGTCACACCTTTAACTTT
     TAATAACGTGTATTATTATAACACGATATTTTAATTCTTTAATTAGCGTTT
     TACATAAAAATAGTAGGCAATATCAAAAATATGCTTTATGAATGATTGTGC
 10  CTACTATAAAATATATTATTTATCATTTTTGTGTAACTCTTGAATCATGCG
     TTCTATCTTATTACCGTATTCGATAGATTCATCATATTCGAATGTTAACTC
     AGGTATAATTCTTAGGCGCATTCGAGAACCAAGTTCAGATTTTATAAACCC
     AGTTGCTTTATGCAAAGCTTTAAACGTATTATCAACTTCTTTATCATTCCC
     TAACACTGTTAAATATACCTTTGCTTGTGAAAGGTCATTGGTTAGTTCAAC
 15  ATCAGTAATTGTTAAAAAACCAACTCTAGGGTCTTTAACTTTATTATTAAC
     AATGTCCATGATTTCCTGTTTCATTTGTTCTCCTACACGTTCTGCTCTTAT
     ATTATTCATCATCTCACCTCTCTTTATTTTTACTTTACTCACAAATAAATG
     TGAAATATCAGATTTAATGATTAACTATTGATGTTCATTTTATTAAATTAT
     CAACAACGAATCAATAAAATAATACTAAATGAGAGACCATAACTAAAGTAT
 20  TGATAGCATAACATTAAGCTTTTCCAAATCAAAAATTTATATCTTCAATTT
     GATTGTTGGTTAAGCAAAGTATGTCATACTCTCCTTTAAGATATTAGGTTT
     ATGAAATATCTTAAAGTACTTAACTTAGCATATAGCCTCTCAATAAGTTTC
     AAAAATTATCTTTGAATTTCTACCATTTCAAACGCTTCAATAATGTCTCCT
     TCTTTGAGATCATTATATTTTTCAATTGTAATACCACATTCATAGCCTTGA
 25  GCTACTTCTTTAGCATCATCTTTGAAACGTTTTAATGTGTCAAGTTCACCT
     TCAAATAACACGATACCATCTCTAATTACGCGTACACCAGCGTTACGAGTG
     ATTTTACCTTCAGTCACATAACTACCAGCAATTGTACCAACTTTAGAAACT
     TTAAATGTTTGACGCACTTCAGCTTGTCCAATGACTTGCTCTTCAAATTCT
     GGGTCAAGTAAACCTTTCATAGCTGATTCTATCTCTTCAATAACATTATAG
 30  ATAACTCTGTGTAATCGCATATCTACATTTTCAGCTTCAGCCGCACGTTTC
     GCACCTGCATCTGGGCGTACATTAAAACCAATAATAATACCATTTGATGCA
     TTTGCTAATGTAACATCTGATTCATTGATAGCACCAACAGCTGTATGAATA
     ATTCGTACATTCACACCTTCAACATCTATTTTCATTAGAGATGCGGCCAAT
     GCTTCAACTGAACCTTGTACATCACCTTTAATGATGACATTTAAATCTTTC
 35  ATTTCACCTTGTTTCATTTGCTCAAATAAATTGTCTAATGAAACATTTTTA
     CTTTCTTGACGTTGCTGTATGACACTTGCCTCATGACGTGCTTCACCAATT
     CGACGTGCTTGTTTTTCATCACCAAATACAACAAAACGATCACCTGCAAGT
     GGAACATCGTTAATACCAGTAATTTCTACAGGTGTTGAAGGACCGGCAGAT
     TTAATTCTTTTTTCCTAAATCATTAACCATTGCACGTATACGTCCATAAGTA
 40  TTACCTACAACAATTGCATCTCCAACGTTTAAAGTACCATTTTGAACAAGT
     AAAGATGCAGCTGGACCTCGTGATTTATCTAATTCAGCCTCAATCACAGTA
     CCTACAGCTTGTTTATTAGGATTAGCTTTAAGTTCTTGTACCTCCGCTACT
     AAACCGATCATTTCTAATAAATCATCAATACCGTCTCCACTCAATGCAGAT
     AGTGGTACAAAGATTGTGTCACCGCCCCAGTCTTCTGGAATTAATCCATAC
 45  TCAGTGAGTTCTTGCATAACACGATCAGGGTTAGCAGTTGGTTTATCAATT
     TTGTTTACTGCAACAATCGTAGGTACTTCTGCTTCTTTAGCGTGATTTATA
     GCTTCAATTGTTTGAGGCATCACACCATCATCAGCGGCCACGACTAAAATT
     GTAATATCAGTAACTTGAGCACCACGTGCACGCATAGTCGTAAATGCAGCA
     TGTCCAGGAGTATCTAAGAACGTAATTTTTTTACCTGAATTTTCAATTTGA
 50  TAAGCACCAATATGTTGAGTGATTCCGCCAGCTTCTCCTTCTGTAACTTTA
     GTGTTACGAATAGAATCTAATAAAGTCGTTTTACCATGGTCTACGTGGCCC
     ATGATTGTAACAACTGCTGGACGTTCAATTGCATCAGAATCATCAGTCTCA
     TCATCAAAATAAATTGATAAATCTTCTTCATCAACGACTACTTCTTTCTCT
     ATTTCAACGCCATAGTCATCTGCAATTAATTCTAATGTTTCTTCATCCAAT
 55  GATTGATTGATATTAGCCATAATACCTAGTAAGAACAATTTTTTAATAATA
     CCAGCTGATTCTACATTTAGCTTTTCAGCTAACTCACCGACAGTTATGCCT
     TCTTGATAAGTGATTTTAGAGGGCATTTCTTTTGTTTCTGCTACCTCATTT
     TGAGGTTTATTATTTTTATTATTCTTTTTATTTTTATTGTTTTTTTGATTC
     TTATTAGTTTTATTGTTTTTATTTTGCTGTTTTCCTTTGTTGTTTTGTTCT
```

```
TTTTTCTTCGTTGGTTTACTATTATTCTTACTTTGTTGTTTTTCTTTATCG
TTAGAATTTTGTTTATTATTAGATTTTTGGTGATTATTTTGAGTATTTTGT
TTATTAGTGTCTTTCGCT
```

5

```
Sequence 3397
step.1002e06.cons.ok
AAAATTCAATATACGTTTCTTTTATAATTTTAGTAATTTCTCTTTAAAAAT
ATCACGTAAAAATTCAGGCATAATATAATCTGTAGATAACCCCTTTAGCAA
TGGGAAATATCTTCCAAATGTATACATATCAACAGTACCTAACGTATATTG
TTTATCATTTTGAATCTCTTCATCGCCTATAAAATAACGGTCTTCTGACTC
AATAAAGCCTAGATTATATAAAATATATCCTCCAAATATATCCCCTCTAAA
ACCAAGACCTTGTGCATGTTCGTGCATATATTCTTGCTTTTGGCTTTTTTG
AATCACTTGCTTTAATTGTTTACCGGTTAATCGAACTCTTACAATATTGAT
TGGATGGGGTAACATGCGATGTATATCATATTCCGTCACTTTATCAGCTTC
AATGCCATTAACTATAAGTCCAGCATTTACGATTGCACAATCAGCCCTTGA
AAACTCATATACACTTTCAGCCAGTAAATACGATGTTCTTGTAACAACATC
TGTTCTTTTGACTAAGTTCACATGATGATTAACTACTGGTTTACTTAGAAG
TGCTCTTCCTTCTTCTTCAAAATGTGTCTCGACTAAGGGAAGTGTTTCAAT
AGGATGAATTTTGGCGATTTTATCAACGATTTTTCCATTTTCAATCGTAAT
ATTAACTTCACCTAAATAATAGCCATATTTTCCGGCAGCTGCCATCAAAAC
ACCATTGTTTATTTCTCCATGTTCAAAATGATGATGCGTATGACTACCAAA
GATAACATCTATTTCCGGAATCTCTTGGCATAACTTTTCATCAAAAAAGAT
ACCGACATGGCTCATAACCATTAAAAGATCATATTCACCTTGATGTGCATT
GATTTCATCTTTGATTGCCGCTAATGGGTCAGTAACAATCCAATCCAGTGC
TCGATAAAAAGGTGTGAACGGTGCCGTTGCTGCAACAAATAAAATACGTGT
TCCTTTTATTTCTTTGATATACGAAGAGGTAATATGATGTGGAAGATGTCC
CTCTTCATCTATGACATTCGTGCAAATCACTTTAAAATCCGCGTCGTTATA
TAGATTTTGTAAAGCATCATGAGAAATTGTCATTCCTTCATTATTTCCAAT
GGTTGCAATATCACAATGTGCTTCATTTAAAAGTTCTATATTTTTATGTCC
TACCGTAGCTTCTGTCACAGGTGCTGATAAATCAACATGGTCACCTATATC
TATATAGAGTGAGGGATGTTCAAGTTGCGGTCTATGTTTTGCCATATAAGC
TTGAATACGAGCATATTCATTTAAATGACTATGAATATCATTCGTGTGATA
TATTGTTAATTCCATAACATGATCCCCCTCTTTAAAATTTATATACTTAAT
TGTACTTTAGCATTAAAGAACAAAGTCTCAAAGCATAACTACTTCGCATTT
AAATAAAAGATTTAATGATTAAATATACACCCATGATAAGCATTACTGTTC
TCAATAACATTACAACTGTGTCGGATTTCATAGATCGATTGACCCTCACAC
CTATTTGTGCACCTATAACACTTGAAATAATGAGAATGATAGAATAGCCCC
AAGCCACATGTCCTTGAAAGATGTGCCCTATTGAACTCATCACACTTGAAA
AGAAAATCATCATCATACTTGTGCCTACTGCAACATGTGGTGGAAATCTAA
AGACGATGAGCATAAGAGGGGGTCATCAATGCACCTCCACCTATGCCAAACA
GTCCTGTTAATAGCCCAATAAATAACGTTGTTATAAAAGCAAACAGGGGTG
GTACACTATAACGGTATGTTTTCCCTTCAGCATCAACATATGTACGCGCAT
ATTTAGGTTTGTCAAAAATTTTAAAAGGTTTGATTTTGTGACGAATCATCA
GTAATATCGACACAAATATCATAAACACGCCAAAGTATAAGTTAAACGAAT
CAAGTGTAAGGTATCTACTTAAAAAAGAACCAACCAAAGATCCAGGTAGTA
GTCCAAATAGAAATATCGAACCATTTTTTATATCTACTTGTTTGGTTTTTA
AATATCCTAATGAGGAAGATAATCCAGTAACAATCAGAATAACTGAAGAAG
TACCTATTGCAATTTGCGTCGTAATACCATGTAGAAGATGATGGTTAACGC
CTAAATAAACAAGTGTGGGGACGATGATGATTCCACCACCAATACCAACGA
TAGATCCTAAAATCGCTGATAATCCCCCTATTAATATTAATAGAATAATTG
TTAATAACATGACATTCGCCTCTTAAAATAGTTTTAATTGTTGCGGTGCTA
AACCTTCATAATCAATGTCTAAAATATTTTGATACTTTTTAGCATTATTAG
CTGCATGACCGCCAGAGTTATTATTAAAAAATTACATATACTTTCTTAGCCT
TTTGATTAAGTATTTCGACTTTACGAGCCAAGTCAGCTAACTCATCATCGC
TATAATCATATAAATATCTTACATCTCGCCATTCTTGATCAGTCATATCTT
TTTTAGTCCAACCATAATGATTACGTCCATGATAACGTACAAAAGCAATTT
CACTAGTAATCCTATTTACTAAAGGAACGCTCCCCTCTTTAACTTGAGGTT
CATCTACCACTGCATGAATGATTTGATGTTGTGTTAAGAAGGATAAAGTTT
```

GTTCTTTATACTGATTGTCAAACCATGATTGATGTCTAAATTCAATGCTCA
TCGGAATATCAGTTAATTGTTTTCTCACATATAAGATGTAATTTATATTTT
GGACGTTACAATCAAACCAAGGTGGAAATTGTACGAGCACCATCGCCAATT
TATTATGCTCTTTTAACGGATTGATCATTTCTTTAAACTGTTTAAATAATA
5   TATCTCTCGACTCTGCATAATCTCTATAGTCAGCATGCAATGTTAATGCTT
GATGAATTTTAACAACGAATTGAAATCTATTCGGCGTTTCATTTATCCATT
TAATTATATTTCGCTCTGGCTGTATCGCATAATACGAAGCATCCAACTCTA
CTATTGGGAAGTGACTAGCATACGTTTTTAACTTATCTGTTTTACGCTCTA
AATCCTCATATAATGTATCATGGTCACCCCAGCCTTCTAAAAACTCTGGAT
10 AAAAATCACAATTACCTTTCACACGCTGATACAAACTAAGTTCTGTGTCAT
TATAACTAAATTCAGAATCCCCTAAATGTAAAAATTTATCTGCATCTTTAT
GTAGTTCATAAATATGATATAAGATACCTGTTTCTTTGTGATTATCACTAA
GAATAATCCACTTAGACATGTTGTTCACCCTCTAAATATTCTTTTAACAGT
CGAATAGCATTTCCACGATGACTGATTTTACCCTTTTCATCATTTGTAATT
15 TCCGCCATCGTTTTATTTAATTCTGGGACAAAAAAATATTGGATCATAACCA
AAACCATTCTTTCCATGTCGTTCAGTTGTTATCACTCCAGAAACTGTTC

Sequence 3398
20 step.1002e07.cons.ok
TTTGGAAGTTCTGTTGATAAAATCAAATTATTCGCTGAATTGACTTCTAAA
TATAAAGATTACCATCCTGAGTATCTGGGTAAGTTTAACAGATATGAAATA
GATAAATATATTCAAGAAATTAGCAAAAGTTATAAAAAAAGAGGTAATGAA
AATGGCTGAAAATCAAGCTAAAAAAGAACAAGTTCATACAGTATTCCAAAA
25 TATTTCTCAAAAATACGATCGACTCAACAACATCATAAGTTTCGAACAGCA
TAAAGTATGGAGAAACATGTTATGAGTACTATGAATGTACAAAAAGGCAG
TAAAGCTTTAGATGTGTGTTGTGGTACTGCAGATTGGACTATTGCATTAAG
CGAAGCAGTAGGCTCAAAAGGCCAAGTAACTGGTCTTGATTTTAGTGAAAA
TATGCTTGAGGTAGGTAAACAGAAAACTGCTTCATTGGAAAATATCCAGCT
30 TGTTCATGGTGATGCTATGAATTTACCTTTTGATGATAACTCATTTGATTA
TGTTACTATTGGTTTTGGCTTACGCAATGTTCCGGATTATTTGTCAGCACT
CAAAGAAATGCATCGAGTTTTAAAACCAGGAGGCATGGTGGTATGTCTTGA
AACAAGCCAACCGACATTGCCATTATTCAAACAAATCTATAGTTTATACTT
TAAATTTGTGATGCCAATTTTTGGTAAAATGTTTGCTAAGTCCAAAGAAGA
35 ATATGAATGGTTGCAACAATCAACATTTAATTTTCCTGACAAACAAACATT
GAAAGGCTTATTCTTCGAAGCAGGATTTAACGATATCATTGTACGTAGCTT
TACAGGTGGCGTTGCTGCAATGCACCTTGGCTATAAAGAAAACAGTTCTAG
CAAAGGTGATTAACGTGGCAAAGTTAAACATTAACAACGAAATAAAGAAAG
TAGAAAAGCGACTTGAAGAAGCAATTATAAGTTCTGATCAAACATTACAAG
40 AAGCCTCATTCCATTTACTATCTTCAGGGGGAAAAAGAGTTAGACCCGCTT
TTGTTATTTTAAGTGGTCAATTTGGCTCTAACAACAAACCTTCAGAAGACA
CGTATCGTGTAGCAGTAGCTTTAGAACTAATTCACATGGCTACCTTAGTCC
ACGATGATGTGATAGATAAAAGTGATAAACGTAGAGGCCGACTCACTATTT
CAAAAAAATGGGACCAAAGTACAGCTATTTTAACAGGAAATTTCTTACTTG
45 CTATGGGGCTCAAGCATTTATCTGAAATCAGTGATACTCGTGTCCATTCGA
CCATTTCTAAATCAATTGTTGATGTGTGTAGAGGAGAACTATTCCAATTTC
AAGATCAATTTAATAGCAATCAAACAATTACTAATTACTTACGTCGTATCA
ACCGTAAAACAGCACTTCTTATTCAACTGTCTACACAAGTTGGTGCGATTA
CTTCCAATGCGTCAAATGACGTTATTCGTAAATTAAAAATGATCGGACATT
50 ATATAGGTATGAGTTTCCAAATAATAGATGATGTGCTAGATTTTACTAGTT
CTGAAAAGAAACTTGGTAAGCCGGTTGGTAGTGACCTTATGAATGGTCATA
TTACATTACCTGTACTATTAGAAATGCGAAAAAATAAGACTTTTAAAGATA
AAATTTCACAACTTAATCCTGACAGTCCTCAACATGCCTTTGAAACTTGTA
TAACAATAATTAGACAGTCCGAAAGCATAGAACAATCAAAACAAATAAGTG
55 AAAAGTATTTAAATAAAGCAATCAATTTAATCGATGAATTAGAGGATGGTC
CTAATAAAGAACTATTTAGAAAGCTTATTAAAAAAAATGGGAAGTCGAAATA
AGTAAGTTTTGTGACAAAAGTTGAAAGCGCTTCATTATCCTGTTACCATAC
TAAGTAGCAGGATATTTTTATAGCTTGCAATGTTTTAAGTTTTTTACATAC
TTATAAAAATGAATGAAACACATGCACAGGGGGACACACGAAGTGGAACGC

```
ACATTTTTGATGATTAAACCGGACGCTGTCCAAAGAAATTTAATTGGTGAG
ATTATATCTCGCATTGAAAAGAAAGGTCTTAAGCTTGTTGGCGGTAAGTTT
ATGCAAGTGCCTATGGAACTAGCTGAAAAACATTATAGTGAACATGAAGGA
AAACCGTTCTACGACAAGTTAATTTCATTCATTACTTCTGCACCCGTTTTT
```
5
```
GCTATGGTAGTAGAAGGTGAAAATGCTGTAGCTGTATCTCGTAAAATTATT
GGAAGCACAAATCCAAGCGAAGCAGCTCCTGGTACAATCCGAGGCGACTAT
GGTTTAAATCTAGGTCGTAATATTATACATGGTTCAGATTCAACAGAATCA
GCACAACGTGAAGTTAAATTATGGTTTACAAGTAGCGAAATTGCTGATTAT
AAAGAACCAAGAGAAGATTGGCTATACGAATAATAAAAATATCAAATAATG
```
10
```
AAATTCATTTTATCCATTAATGTCTGTGACGCATATTTGTGTCTCAGGCAT
TTTTTATGATTAACCAAATTTTATTTGCTTAAAATAATACTTATTGAATTA
GTCTATTTATTTTTTTATCTAGGGATTTAACGATGCAATACCGTATATTGT
TCGATTATATAGCAATTTTATCTTTTGACAATTGAGCATTCATTATTAAGT
AATGACCTACTTTTAATATATCAATCATGTATCTAATGATTTCGGTCTTTT
```
15
```
GAGTTTTTTAATAGACTGATCTTTAATTGGTATAATGTCACTATGACATAA
CAATCATCTTGGTGCTATTCTTTTTTAAATATTTATAACTGCTAAAATAAA
AATCTCCTTGTTTGTCTAACTGACTTATGTTGTGTGATTGAATATGATAAT
ATATTAAACATATTACAATTTAATGTATTGAAGAATAGAAAGGGGCGTGCG
ATCATGAGATACCTTACATCCGGTGAATCACATGGACCACAACTTACAGTC
```
20
```
ATCATAGAAGCTTTAGAGCACAGTGGCGA
```


Sequence 3399
step.1002e08.cons.ok
25
```
TTTAAATTGTCAGATGATCAAGAAAATGAAATTGAGCAATTATTAGATCAA
ACCAATCCTGATTTACCACGACCAGTAGGAGAGGATATTGTACATTATTCA
GATTATTTTGAAGGTGCACAAAAGTATCTAAGTTATCTTAAATCAACTGTT
GATGTTAATTTTGAGGGTCTTAAAATTGTATTAGATGGTGCAAACGGGTCA
ACTTCTTCTTTAGCCCCATTCTTGTTTGGCGATTTAGAAGCGGATACTGAG
```
30
```
ACAATTGGATGTAATCCAGATGGTTATAACATTAATGAACAATGTGGCTCT
ACTCATCCAGAAAAATTAGCTGAAGCTGTGTTAGAAACTGAAAGTGACTTT
GGTTTAGCTTTTGATGGAGATGGCGATCGAATTATTGCGGTAGATGAAAAT
GGACAAATTGTAGATGGAGATCAAATTATGTTCATTATTGGTCAAGAGATG
TATAAAAACCAAGAACTCAATGGAAATATGATAGTTTCGACAGTAATGAGT
```
35
```
AACCTTGGTTTCTACAAAGCTCTAGAAAAAGAAGGTATTCAGTCAAACAAA
ACTAAAGTTGGAGATCGCTATGTTGTCGAGGAAATGAGAAGAGGAAATTAT
AATCTTGGTGGTGAACAATCCGGTCATATCGTATTAATGGATTACAATACT
ACTGGTGATGGATTATTAACGGGTGTTCAGTTGGCTTCCGTTATTAAAATG
AGTGGTAAAACTCTAAGCGAGTTAGCTTCTCAAATGAAAAAGTACCCACAA
```
40
```
TCTTTAATTAATGTGAGAGTGACTGACAAATATCGTGTTGAAGAGAATATT
CATGTTCAAGAGATAATGACGAAAGTTGAAACAGAGATGAATGGTGAAGGA
AGAATTCTTGTTCGTCCTTCTGGAACTGAACCTTTAGTACGTGTAATGGTT
GAGGCTGCAACTGACGCGGATGCTGAAAGATATGCTCAAAGTATCGCTGAC
GTTGTTGAAGACAAAATGGGCTTAGATAAATAATACTTTCATGTTTAAATT
```
45
```
CATATATAAAACAGATTTATGATGTTACAGACATAAATCTGTTTTTTTTAA
CAGAATTTTATCTTTTAAAAAGATTTGAAAAATATATTTTTTTTATTTTGAG
AATAATTCATCGATGTATCTATTCAGAACTATTACTATTTGTGAAAAATAC
TGTTAAATCGTTTGCAAATTAATTCGTTGACAGTGTACAATGAATATTAAT
GGAAAGGAGGAGAGTCTAGGTTATTAATAATTAAAGCGCCTGTGCAAAATA
```
50
```
CCTAAATAGTTATCAAGTATTAAAAGTTAAAATAGTAGAACTTTTATAGAT
AACGATTTGGATTTTGTAGACGAGGAGGATAGTGATCGAATCAGATCGGCG
GATGCTATCCCGGATGTGGCACATTCGTTAGCTTATTAAGTAAATCATTAA
GGTGACTTAGTGGACAAAGTTAATAAGATCGCCAAACTAAATAAATTATTT
TGGGAAACTACGGATAGCGTGTATAAAATTCATATAATAGCACGCTTATTT
```
55
```
CATGATGTATTCGTTGATAATGATTATGATAGGGGATTATCAAATTTAAGC
AGTGATAATTAATCATAATAGCCGTGTTATCCCATTTGTTGGAGGATATTA
TATGTGTGGAATTGTTGGTTATATTGGCTACGATAATGCCAAAGAATTACT
ATTAAAAGGGTTAGAAAAATTAGAATATCGTGGTTATGACTCAGCAGGTAT
TGCTGTAGTTAATGATGATGGTACAAAAACTTTTTAAAGAAAAAGGAAGAAT
```
```
911
```

```
TGCTGAATTACGTAAAGTTGCAGATAATAGTGATGAGGATGGTACGTTAGG
AATTGGTCATACACGTTGGGCGACACATGGTGTTCCAAATTATGAAAATTC
ACATCCACACCAGTCAACATCTGGACGTTTTACATTAGTTCATAATGGTGT
AATTGAAAATTATGAAGAATTAAAAGCTGAATATTTATCTGATGTCACTTT
TTCATCAGAAACTGATACGGAAGTTATTGTACAATTAGTAGATTATTTTTC
TAGACAAGGATTAGCTACAGAAGATGCATTTACAAAAGTAGTTAAATTATT
ACATGGTTCATATGCTTTAGGATTATTAGATGATAATGATAAAGATACTAT
TTATGTGGCTAAAAACAAGTCTCCGCTTTTAGTAGGTGTAGGTGAAGGTTT
CAATGTTATTGCTTCTGATGCTCTAGCAATGTTACAAACTACAAACCAATA
CAAAGAGATACATGACCATGAAATAGTTATTGTTAAGCGAGACACAGTAGA
AATTAAAGATCTTGAGGGGCACATTCAACAACGTGATACGTATACGGCAGA
AATAGATGCTGCTGATGCAGAAAAAGGCGTATATGATCATTACATGTTAAA
AGAAATTCATGAACAGCCTGCAGTGATGCGTCGCATTATTCAAGAATATCA
AGATGAAAAAGGTAATTTAAAAATCGATTCAGAGATTATTAATGATGTAGC
AGATGCTGATCGTATTTACATCGTTGCAGCTGGTACTAGTTATCATGCTGG
ATTGGTTGGTAAAGAATTTATTGAAAAATGGGCAGGTGTACCTACTGAGGT
TCATGTAGCTTCTGAATTTGTATATAATATGCCACTTCTTTCTGAAAAACC
ACTATTTATTTATATTTCACAATCTGGTGAAACAGCTGATAGTCGTGCTGT
ATTAGTTGAAACAAATAAGTTAGGTCACAAATCATTAACAATTACTAATGT
TGCTGGTTCAACATTATCACGTGAAGCGGATCATACATTACTTTTACATGC
TGGACCTGAGATTGCAGTCGCATCTACAAAAGCATATACAGCGCAAATTGC
TGTTTTATCTATCTTATCTCAAATTGTTGCTAAAAATCATGGTCGTGAAAC
CGATGTTGATTTATTAAGAGAACTAGCTAAGGTTACTACAGCTATTGAAAC
AATTGTTGACGATGCACCTAAGATGGAGCAAATTGCAACGGATTTCTTAAA
AACTACTCGTAATGCATTCTTCATTGGACGAACAATTGATTATAATGTTAG
TTTAGAAGGTGCATTAAAATTAAAAGAAATTTCTTATATTCAAGCTGAAGG
ATTTGCAGGTGGGGAATTAAAGCACGGAACAATCGCTTTGATTGAAGATGG
CACACCTGTTATAGGTTTAGCTACACAAGAAACGTTAATCTATCAATTCG
TGGAAATATGAAAGAAGTACTTTAGAGCACAGTGGCGATGATATCA
```

Sequence 3400
step.1002e10.cons.ok
```
TTCCTCATAACGACCCTTTTCAACTTCTCTTGGATCATATTTTGGTTTCAT
TTCCATATGTCTACCTCCTAAAAAATAAAAGAACATCCATCCTATAAAAA
ATAGGACGGATGTTCTGTTCCGTGGTACCACCTATATTCAAGAAGAAATAA
TTCATTCTTTTCTTCAAGCACTTAAGTCTCTGATTAACGCTCATACACGCT
TCTACCTACTTGTCGTCTATGTTTCAATAGAAGTTATAAAATAGGGCTACC
TTCAGACTTAATCATTTACATATTCACAGCCACCATATGCTCTCTTTAAAA
TGATATAAACTTACTCTTCCCAAATAATATATTCTTTTACTCATTTATATT
ATAATGGATATTAAAATAAGTCAATGTATTGGAGATGAAATTATGAATGAA
TGTGCTTTTAACACAACTGATCCTATCTATATAGAGTATCATGATTATTAT
TGGGGACAACCAATCTATGATAGCAAAGAATTATTTAAGCTAATGGCTCTT
GAATCACAACATGCAGGATTATCATGGTTGACAATATTAAAGAAAAAGAA
TCGTATGAACAGGCATTTTACAACTTTGAACCTCAATTTATCGCACATATG
ACTGAACAGGATATTGATTATTTAATGAAATTCCCAAACATTATTCATAAT
CGTAAAAAGTTAGAAGCAATTGTGAGCCAAGCGAAGGGATATTTAAAAATT
GAAAAAGACTATGGTAGTTTTAGTAAATTTTTATGGTCTTATGTAAATCAT
CAACCCATAAATATGGGTTATAAAAAAACCTAGAGATCGTAAAAAAGTTGAT
CAAAGAGCTACTCAATTATCAAAAGATTTAAAAGCATATGGTTTTAAATTT
TTAGGTCCAGTAACCGTGTTCTCATTCTTAGAAGCTGCTGGATTGTATGAC
TCACATCTTGAAGGATGTCCTTTCAAACCAAATCATGAGTGAAGCATAGAC
ATAAGTTGCGCTGCGGATATTCAAACAATATGAGTTTTTCACGCATACTAA
AAGTTTACAACGTTTAATTTATCTAGATTTTACTTTTAAGTAATTACAATT
ACTCGCAAATATTTTCGAAGTTTTATTTAAAAATGTATAACTTAAGTCATT
AAAATCATTTATTATTTAATTTAACATTAATAAAATAACCAATTAGTGGCG
CAATGACAAATAATATAAAAAAATATTCTAAAAATGTGATAGCTTGAAATCA
TCGCTACATCAGCTCCAGTAGCCATAGCCACTAAAACTATTTGACTCATAC
CTCCTGGTGCTGCTCCTAAAAACAATTCATTGATGGATTCATTAGTAATCA
```
```
5
10
15
20
25
30
35
40
45
50
55
```

```
      AATGTATTCCTATTATCATTATAAACGTTGTGACTATGAGCATTATATTTT
      GAAAGGCTATTGCTATTGCAATTCTTCCCTTTAAATCACTCATTAAGTTGG
      CAATCTGTAATCCAATACGTATCATATAAATAAGTTGCGCTGTGGCTAACA
      ACCAATGATCTAGTGAAAATGTTAAATGTGTTGTCATATTCCATATAATTA
  5   AAACTATTATAGGTGCTAATAATTGTTTAGTGGGGAAGTTAATTTTTGACA
      TTCCTATATAGATGATTCCCACCATTGAGAATAAGATGATTATTTGCCATA
      TATTTAAAGTCTGAGAAAGTGTGGGTACTTCCATTGTAGTATGATTCATTT
      CATGATGGTTATCCTGAAAAAAATACGAAATAAGTGGTACTAAAATAACAA
      CAAATATTACACGTGATGTCTGTGTTAAACTCACAACTAATATATTTGCTT
 10   TCTTATTTTCTTCTGCCATCACTAACATTTGGCTTAGCGCACCTGGTATAA
      CACTTAAAATTGCAGTTTCTAAATTTACTTGTGCAATTTTCTTAAAAAAGA
      ATGCAATTATCAAAGCTAATAAAATTAGTAGGATAGTGACAAAAACGATAG
      TTAGCCAATTTTTACTTATGTCTTTAATCACTTGTTGTGTGAAGGTAGAGC
      CAATTTGAACTCCAAGTAGTATTAAACCAATTTGACTCAACCAAAATGGCC
 15   ATCGTATTTTTAATTTTAATACTTTTACACATAATAACGCCGCTAATATTG
      GTCCAAACATAAATGGCAAAATCACATGTAACATTTTTAATAACAGACTTA
      ACATAATAGCAATTAATAATATGATTAAATTATTAATAAAAATTCCGCTTCA
      AAATACCACTCTACTTTCGTTAAAAGTAAGTCACAAAATTTCTAACTTTAT
      TTAAATTCAAAGAGTAGAACAAAATGCATCTTAAATTCCTAGTCCACCATC
 20   GAACCCCAACTTGCCCTGTCTGTTGAATTTCTCAGTGAAATTCTTTTTGTT
      AGGGCCCCAAACCCCAACTCGCCTTGTCTGTTGAATTTCTAAGTGGAATTT
      ACTGTATTGATGACATAACATAGCTTTAAAAGTTGATATAAGTGCATTTTT
      AATTCAGTTATAAACTACTTATAAGTTAAGAAAGTCTAGGACATTGACTTT
      TGCCCTAGACTCAATTAATTTAAATAATTATATTTTAAAATTATCACACAA
 25   TACGACTTAATGCATTATCAAATGCTTGGATAGTTTTCTCAATATCATCTT
      TAGTATGTGCAGTTGATAAAAATGTTCCTTCAAATTGTGAAGGTGGTAGGA
      AAACACCTTCCTTAGCCATTTCTCTATACATATTACTAAATAATTTTAAAT
      CACTTTTATTTGCTTCCTCAAAATTTGTTACAGGCCCCTCATTTAAGAAGT
      AACCAATCATTGAACCAGCGCGATTTACTGTGATTGGAACATTATGCTTAG
 30   CAAATACCTCTTTTAATCCTTTTTCAAGTATATCTCCTAGAGAATTAAAAT
      ACTCATAAGATTCAGGAGTAAGTTGACTCAATGTTTCATAACCACTAGTCA
      TTGCTAAAGGATTACCTGAAAGTGTGCCAGCTTGATAAATAGTCCCAACAG
      GAGCAATGTAATCCATAATTTCTTTTTTACCACCAAAAGCTCCAACGGGTA
      AACCTCCACCTATCACTTTTCCTAAGCAAGTTAAATCAGGTGTTACACCAA
 35   AGTATCCTTGCGCACAATTATAACCTACACGGAAACCAGTCATCACTTCAT
      CAAATATAAGTAATGCTCCATATTCATTAGTAATATCTCTTAAACCTTGTA
      GAAATCCATTCACTGGAGGCACTACACCCATATTTCCAGCAACCGGTTCAA
      CAATAACACCAGCAATATCATCGCCATATTTTTCGAACGCTAATTTAAGTG
      AATCTAAATCATTATATGGCACCGTGATAGTGTTTTTAGCAATACCTTCAG
 40   GGACGCCTGGTGAATCAGGTAAACCTAGTGTTGCAACACCTGATCCTGCTT
      TAATCAATAAAGAATCACTGTGTCCATGATAACACCCTTCAAATTTTATAA
      TTTTATTACGTCCTGTATAACCCCTAGCTAAACGAAGTGTGTCTAAAGTAG
      CTTCAGTTCCTGAGGAAACCATTCTTACTTTTTCAATTGAAGGTACACGGT
      CAATCACAAGTTCAGCAAGTTTATTTTCTTGAAGTGTTGAAGCGCCGAAGC
 45   TTGTACCTTTATCTACTGCTTCATGTAATTTGGATATAACTTGTTGATTTT
      TATGTCCCAGAATTAATGGGCCCCAACTTAGCACATAATCAATGTATTCAT
      TTCCATCAATATCATATATTTTAGATCCTTCACCATGATCCATAAAAATAG
      CTGGTGTGTCTACTGATTTAAATGCTCTTACGGGACTGTTAACACCGCCAG
      GCATTAATTTCTCAGCTTGCTCCATTGCTTTAATAGATTTTTCAAAACTCA
 50   TCATATGACCTCCTCAATTATCATCTATTTATCTAAATAACGACAGATATC
      TTTTGCAAAATAAGTAATTATTAAATCAGCACCTGCACGTTTCATAGATAT
      CATTTGTTCCATAACAATTTTCTCTTCATCTATCCAACCATTTAACGCTGC
      TGCTTTTGTCATACTATATTCTCCACTAACGTTG
```

55

```
Sequence 3401
step.1002e12.cons.ok
GGTAGGTGTATTATTTATATTTACTACTTTGAATATGGCTTCAATTTCTGA
TAGAAAAGTACATGCCAAAATTAAAAATATAAGAAGATAGGTGAGATAAAG
```

```
TGGAAACAATAGGAGTATCATATATTTAAAAGAAGGTTCCCAAAAGTTAAT
GATTATCAATCGAGGACCTATTGTAGATATAGATAATCAAAAATATATATT
TGATTATTCAGCATGTAATTATCCTGTCGGTGTCGTAGAAGATCAAATATA
CTATTTTAACGAAGACAATATCGATAAAGTCGTTTTTGAAGGTTATTCTGA
TCAAGATGAGATGAGATTTCAAGAATTATTTAAAGAAATGAAAAATAATTT
AGATGATGATATTCAACAAGGTATTGTTCAAAAGCAAGATAATTTAGGTTT
AATTTAACATAAGTCATGATTATCTGAAGTTATCTATTATCGAATATTCAG
AGATGATATAATATATGAGTCGCCTATCTCTCAGGCGTCAATTGACGAAGA
GAGGAGGTGCATTAATTGCTAATCATTTTCGTTCACATCACGACCACAGTC
CTCAGTTGTTGTATTATTGCATTATTTACTCATTGGATATGTAATTGCAAT
AATAAATAGGCAACCATAGACACACATAAAAATCCCCACACTATTGCAGTA
GTGTGGGGATTGGTGCATTAATTGCTAATCATTTTCATTATCTAGATTATA
ACACTGGTTAGTAGAGAAATGCAAAGATAGACAAATAATTAATGACTGAGG
ATAAAGATATAAAACAAACATACTTTCCTCATTTTGATACCATAGCCACCA
ACTTTTTACTTGTGTTCTATTTCTAATATTTAATTCTTTCATAATTTCTTT
TGTTGAAAATTCTGCTGCTTTCATTTCAACGGCTTTATACTTTGTTTCTAC
TGAATAAGAAGCTCTTTATTTTGGAGCAACAGGTGTTTTAAGCTATCGAAA
ATTTTTCCCATTTATTTATAATTTGTTTCAAAAAGGATACTTTGAAAATTC
ACTAACAATTATTGGCAGTGGTTTAAATGAGTTAACTACTGATGAATTTAG
AGAAAAAGTCAAAAATGCTATTCAAAATAATATTGAAAACTCAAAGGAAAT
TGGTGCGTTTTTAAAACGTTTATTTTATAAACAACAGGACGCTAATAGTAA
AGATAGCTACCAAAAATTTTTGGAAATGAGTTTAGAACTAGATGATAAGTT
TGACCTAAAAGAAAATCGATTGTTCTATCTTGCAATGTCCCCTAAATTTTT
TGGAGTTGCAACAAACCACTTGAAAGAGTCTGGATTAACGAACGTAAAAGG
GGTGATGCGAATAATTATAGAAAAACCATTTGGTGATGATTTAAAATCTGC
AAAAAATTAAATAATCAAATAAGAAAGTCTTTTAAGAAGAAAAAATATTTG
GAATTGATCACTATTTAGGTAAAGAAATGATTCAAAATATTGAACGTCTAC
GATTTTGAAACACTATATTTGAACCACTTTGGAATAATAAGTATATTTCCA
ATATACAAGCAACTTCTTCAGAAACGATTAGTATAGAATATCGTGGTGGTT
ACTACGAATCAAGCGGCGCCCTAAAATTCATGGTTCAAAATCATTTATTAC
AAATGGTTTAATTACTTGCAATAGAACCTCCAATTAGTAGAAAAAGTAGTG
ATATAAGAAAAGAAACATTTCAAGTCTTAAAATCCTTGAAATGTTTTAAT
CCAAATGAAATTAAAGAAAGTTTCGTTCGTGGTCAATATGATGGGGGAATG
ATGAATAATGAGTTCGTTCCTGCATATAGAAATGAACCTAATGTAAATTCA
CAATCTAATACTGAAACTTTTGTAGCAGGTAAAATAGAAATTGAAAACTCT
AAATGGGCTAGTGTTACATTTTATATTCGTACAGAAAAAAGAATGAAAAAA
ATCTATCCAAATCGTTATAGAGTTTAAAAAAATCCAATGCAATTTTATTATA
GTAATTGTGAACAGAATGCATCAAACTTATTAGTCATCAACGTACAACCTA
ATGAAGGATTTTCTTTATGTGTGAATGGTAAGAAAAGTAATCAAAATAATG
AAATGCAAAAAGTGAAGCTTTCTTATACTATGCCGATTAAAGATAAAATGA
ACACAGTTGATGCATATGAAAATCTTATTTACGATACATTAATTGGAGAAC
AAACAAAATTTACGCATTGGGAAGAATTAAAAATTCTTGGAAATTTATTGA
TGATATTGAAAATGTATGGAAACAAGAATAGCCACAGTTTCCTAATTATGC
CTTTGGATGCTATGGGCCTAAAGAAAGTGAAAAATTACTTAGTGAAGACGG
ATTGAATTGGTGGAATAATTTGTAATATAAAAGACCTAAAGTTGTCGTGTT
TTGAATGAGTAGTTAATCCTTAATTAATATGCAATAATCAATGTTATAAAG
TTTATAAAGGGAGGACAATTAATGGATTTTATTAATATTACAGGTGCTTCA
CAAAATAACTTGAAAAACATAGATGTAAATATCCCAAAACACTTAGTAACG
GTATTTACAGGTCGTTCTGGTTCAGGGAAATCATCTTTAGTGTTTAATACT
GTTGCTGCGGAGTCTGAACAGCTACTAAATGAAAGTTATTCTAGTTATATT
CAATTTCATTTAAATCAACAACCCAGACCGAAAGTAAAGAAAATTAAAAAT
CTTCCTGTAGCAATGACGATTAATCAGAAAAGATTCAATGGGAATTCTCGC
TCCACGGTAGGAACAGTTTCAGATATATATGCTTCTGTTAGATTACTGTGG
TCTAGAATAGGCGAACCGTTTGTTGGTTATTCAGATGCATATTCCTTCAAT
AGTCCTAAGGGCATGTGTAAAACTTGTGAGGGATTAGGATATATTGAAGAC
ATTAACTTAGATGAATTGCTAGATTGGGATAAGTCTTTAAATGAAGGTGCA
ATAGACTTTCCTTCTTTTGGACCAGACAAAGAGCGTGGTAAAGCCTATCGA
GATAGT
```

Sequence 3402
step.1002f01.cons.ok
GTATTCAACAAGTTTTTTTAATTTATCTTTTCGTGTTAAATGTCTAAATCC
TTTATCTAAACTTTTCATTTATGTACTCCTTTAAAATATCTTTATATTACT
GTAGTAATGTAACCTTTTAACTTTCGCTTTAAATATCAACTGCTAACATTG
TCTTGCACTCAAATAATTCAAAACTTTTCACAATTTTTGTAAAGCGCTTAC
ATATTTATTATATATTCCTGGTCATGTCTGTTATATTTCATTGTATAAAAC
GAATTTTCTTTGGTATGATACTAGTGCTACTTCAAAAAATGAAAGGATGTC
ACACCTTATGAATATAGGTATAGATAAAATAAGTTTCTATGTACCCAAATA
TTATGTAGACATGGCTAAACTTGCAGAAGCGCGCCAAGTCGATCCTAATAA
ATTTTTAATTGGAATTGGTCAAACTGAAATGACTGTGAGCCCAGTGAATCA
AGATATCGTATCTATGGGAGCCAATGCTGCTAAAGATATTATAACAGAAGA
AGATAAAAAGAATATTGGTATGGTTATAGTAGCAACTGAGTCTGCGATTGA
TAATGCCAAAGCAGCAGCCGTTCAAATTCACCATCTTTTAGGTATTCAACC
CTTTGCAAGATGCTTTGAAATGAAAGAGGCTTGTTATGCAGCAACACCTGC
AATTCAACTTGCCAAAGATTATCTTGCTCAACGCCCTAACGAAAAGGTTCT
TGTCATTGCTAGTGACACAGCTCGTTATGGTATTCATTCTGGTGGTGAGCC
TACTCAAGGTGCCGGTGCAGTTGCAATGATGATTTCACATAACCCAAGTAT
TTTAAAACTTAATGATGATGCCGTAGCATATACTGAAGACGTTTATGATTT
CTGGCGTCCAACGGGTCATCAATATCCCTTAGTTGCTGGTGCATTGTCGAA
AGATGCCTATATCAAGTCATTCCAAGAAAGTTGGAATGAATATGCACGTCG
CCATAATAAAACACTCGCTGATTTCGCTTCACTATGTTTCCATGTACCATT
CACCAAAATGGGACAAAAAGCTTTAGATTCTATTATTAATCATGCCGATGA
AACTACACAAGACCGTCTTAACTCTAGTTACCAAGATGCAGTTGATTATAA
TCGTTATGTCGGTAATATTTACACAGGGTCCTTATATTTAAGTCTCATCTC
TTTATTAGAAACACGTGATTTAAAAGGCGGACAAACGATTGGTCTCTTTAG
TTATGGTTCTGGTTCTGTAGGCGAGTTCTTTAGTGGAACATTAGTAGATGG
ATTCAAGGAGCAATTAGATGTTGAGCGCCACAAATTTTTATTAAATAATAG
AATAGAGGTTTCTGTTGATGAATATGAACATTTCTTCAAACGCTTTGACCA
ATTAGAATTGAATCATGAACTTGAAAAAATCAAATGCAGATCGTGACATTTT
CTATTTAAAATCTATTGATAACAATATTCGTGAATATCATATAGCAGAATA
ATCTAAACTCAATTCTTTTGCTTTTTTAATTTTTTTAAAATGAAAGAAACGA
ATTTACTATTCATACTTGCCCCCTAAAAACAACGTCATATAGCTGTTTTTA
GGGGGTTTTTAATCCCAAGATAGTGCATTTCATCTATTATAGACCTCTACT
ATTTATATACTGTTTACCTACTATTCATAGATAAAAACGAATCCTATCCAT
AGATCAACATTGACAAATAACTAAAGCAATTCTATACAATAACGTATCTCA
AGGTTTCGTTGTTTTTTTTAGGTCTGTATAAATGTGTCATATCTATTTCTTC
CAATTTTAATAATTTTATTTTGTTGTCTAAAGTACCTCCGTATCCTGTTAA
ATTTCCATCTTTTACCAACCACTCGATGACATGGTATAAGTATAGAAATAGG
ATTACTCCCAACAGCACCTCCTACAGCCTGAGCAGACATTGCTGGCTTATT
CATTGCTCGTCCTACTGTTTTAGCAATCTCTCCATATGTTTTAAGTTCTCT
ATAACGAATTGTCTGAAGTTCATTCCATACCTTCTGTTGAAAATCAGTGCC
TTTTGGTGCTAACGGTACTTTTATCTCAGGATAATTCCCTTTAAAATACTC
TCTTAACCACGCTTTAGTATCTTTAAAAAACATCTAACGTATCCTTTGTTTT
AATTCCTTTCAAATGTTGATTAGGGTATAAGACATGTGACAACGAGATGCC
ATCGTTTGTAATTAATTGCAATAAGCCAACGGGCGAGTCATACATTGATTG
ATACATATTTCTCCAACTCCTTAAATGTGTATATCATTCTAGTCATATCCT
TATAAACTAACGCATTCTATTCTGAACATAATTTTATAATATAGATAATTG
CTATATTTATGAAATTGGATAAACATTCGAATAGAATGTGATACACCAAAG
AATTGATACGTTTTATTATCTTTAGCAACCTCTTTTAATTTCTTAGTATCT
TCATGCTTAACATAATTATTTAATTGCTTTTCACTTGAATGACACGCCCAT
GAATCAAAATAGAAAAAGTTTACAAAAATATAACAAATTGCTAAAAAAATT
ATTATACCAATAACATAGAATACTTTTTTTCATTAAAGTTCACCTAGTTTTC
AAAAAAATTTTTATAATTTATGATAACTTAAAAGTGTATATTTTTTCTACA
TAATATATTTAATTTTCAAGGAGATGTAAAAAAGTTGAAAAATTTCGCAAA
ACTAATTTTGGTGGGCATTTTAGTATCGGGGTCAGGGATAGCGAGTGTACA
AACAAATATAACTCACGCAAAAGAAAGTCACGATTCAACTCCTCAAAATAT
TAAATTAGTGGGAACGTATGATACTTCTCAAGTTGATTCCAAAACGATGAA

ACAATTTAAAGAAATAGAAAAAGAAGATAATAATTTCCACAT

Sequence 3403
step.1002f02.cons.ok
CAGAAGGATCAATCACACCATAGCGATGTGTATCAGATTCTGGCACTTCTT
GAACACCTATTACTGAATGGCCTGTTTCTTCATAAACATCCATTAATTGTT
TAATAGCTGGTGTATCAGACTCTACAATGTCATCTCCTAATAACACTGCAA
ATGGTTCGTTACCTATAAACTGTTTTGCAGTATGAATTGCATGTCCTAGCC
CTTTTTGTTCTTTTTGTCGCACGTAAAAAATATTAGCTAAATCTGTTGAAT
ATTGTACTTTTTCAAGTAAATCTGCTTTTCCTTTACTTTCAAGTACTATCT
CTAGTTCTTTTTGATTGTCAAAGTGATCCTCAATTGCACGTTTATGCTTGC
CAGTCACTATAATAATATCTTCTATTCCTGCATTAAAAGCTTCTTCTACAA
TATATTGAATTGTTGGTTTATCTAATATTGGTAACATTTCTTTTGGCATCG
CCTTAGTTGCTGGTAAAAAACGAGTACCTAAACCAGCAGCTGGTATGATTG
CTTTTTTTATGTTTTTCAACGTATAAACTCCTTTAACAACCTAAGTTTTTG
ACTACTTAATAAAGTGTATCACTATTTAAAATGCTTTCATAGTAATGTTTT
TTCACCACAATCTTTTTCAATATTAACATTAACTTAATCAAGAATTGAGTT
AGAAATTAAAAATACTCGAATAGAAATTAACTTTGTTTTAAATATATCGGT
GACAGACATCTACTGAAAAACACTCGCTATATTTTTAAAATGAGAGCTTCA
GGAATTTTTTCTCTATACAAATAAACTGATAAATAAAATAAATAAAATACC
AGATACTGAAATAATTGTCTCTAATAACGACCATGTTAAAAATGTTTCTTT
AACTGTCAGCCCGAAATATTCTTTAAACATCCAGAATCCAGCATCATTGAC
GTGAGAGAGAATTACGCTACCTGCACCTATTGCAAGAACGACCAAAGCTAC
ATTAACATCTGAATGTTCTAAAAGTGGTAACACAATGCCTGTTGTTGATAC
GGCAGCAACTGTAGCTGATCCTAATGAAATCCTTAAGACTGCAGCTACAAT
CCATGCTAATAAAATGGGCGACATGCTTGTTCCTTCAAACATCTTAGCGAT
TGTATCACCTACGCCACCATCGATGAGCACTTGTTTAAATGTACCACCACC
ACCGATGATGAGTAACATCATGCCGATTGGATAAATAGCATGCGTAACTGA
TTTCATAATGTCTTCCATTTTCCGTTGTTGCTTCATTCCCATTGTAAAGAT
TGCAAATATCACTGCGATAAGCATTGCAGTACCAGCAGTTCCTATGAAGTA
AACAATTTGTTCGAATACATTTGTTGCTTCTTCGTGTCCTGTACTATGGAT
GGTGACGGTGCAGCTGCAATGCGTTATACCGAAGCACGTATGACTAAAATA
ACATTAGAACTTTTACGTGATATTAACAAAGACACAATTGATTTTATTGAC
AACTATGATGGTAATGAAAGAGAGCCGTCAGTCTTACCTGCACGTTTCCCT
AACTTACTAGTAAATGGTGCGGCAGGAATTGCCGTAGGTATGGCTACAAAT
ATTCCTCCCCACAATTTAACTGAAGTTATTGATGGTGTGCTCAGTTTAAGT
AAGAATCCAGACATCACAATTAATGAGCTGATGGAAGACATACAAGGTCCT
GATTTTCCTACAGCTGGTTTAGTACTAGGGAAAAGTGGTATTCGTCGAGCT
TATGAAACAGGTCGTGGGTCAATTCAAATGCGTTCTCGTGCTGAAATAGAA
GAACGTGGTGGTGGCCGTCAACGTATTGTCGTAACGGAAATACCTTTCCAA
GTCAATAAAGCGCGTATGATTGAAAAAAATCGCAGAGTTAGTTAGAGATAAG
AAAATCGACGGTATTACAGATTTACGTGATGAAACAAGTTTGCGTACAGGT
GTAAGAGTAGTTATTGATGTACGTAAAGATGCAAATGCGAGTGTTATTTTA
AATAATTTATATAAACAAACGCCATTACAAACATCATTTGGTGTGAATATG
ATTGCTTTAGTGAATGGTAGACCTAAACTAATCAATTTTAAAAGAAGCACTT
ATCCATTACTTAGAACACCAAAAAACAGTGGTTAGACGACGTACTGAATAT
AATCTTAAAAAAGCAAGAGACCGTGCCCATATTCTAGAAGGTTTACGAATA
GCACTAGATCATATTGATGAAATTATCACAACAATTCGTGAATCGGACACT
GATAAAATTGCGATGGCAAGTTTACAAGAGCGTTTTAAACTAACTGAACGT
CAAGCTCAAGCAATTTTAGATATGCGTTTAAGACGTTTAACTGGATTAGAA
AGAGATAAAATAGAATCTGAGTATAATGAACTTCTAGAATATATTAAAGAG
TTAGAAGAGATTTTAGCTGATGAAGAAGTACTATTACAATTAGTTCGTGAT
GAATTGACTGAAATTAAAGAACGTTTCGGCGATGAACGTCGCACTGAAATT
CAATTAGGTGGTCTAGAAGATCTTGAAGATGAAGACTTAATCCCTGAAGAA
CAAATTGTTATTACATTAAGTCATAATAACTATATTAAACGTTTACCAGTA
TCTACATATCGTTCTCAAAATCGTGGTGGTCGTGGCATACAAGGTATGAAC
ACGTTGGATGAGGACTTCGTTAGTCAATTGGTAACAATGAGTACACATGAT
TATGTTCTGTTCTTTACGAATAAAGGTCGTGTATATAAACTCAAAGGTTAT

GAAGTTCCTGAGTTGTCACGTCAATCCAAAGGCATACCTATTATTAATGCG
ATTGAACTCGAAAATGACGAAACAATAAGTACGATGATTGCAGTTAAAGAC
CTTGAAAGTGAAGAAGATTATCTCGTATTTGCGACAAAACAAGGTATCGTT
AAACGTTCATCATTAAGTAACTTCTCCCGTATTAACAAAAACGGTAAAATT
5   GCAATTAACTTTAAAGAAGATGATGAATTAATTGCAG

        Sequence 3404
        step.1002f03.cons.ok
10   GGACAGTATTCCTATTATCACACCCATAATAATGGGACGAATACTGGTTTG
        AACTTGTAATAATGAGATATTCATCATTATTAACGAAACACATATCACTGT
        AAACATATAAACACTCAAAGTTAGAGTTAAATGATTACTTTTAAAGAATAC
        AACATTCATAAGTGGAAGCAGATTAATTAATCCTATACTAGCTGCTGTACT
        TATTACTACCGTGATTTTAATTGATGCATTTGCATAACGATTCATATGAAT
15   TTGATTATGTTCACGAATTGCTTGAGTAAGTAATGGGATAAGAACGAAACT
        AAAAGTCGTAGTTACAATCAAACCCATTTGTATAAATGAAGCACCACGATC
        ATAAATGCCTTTTTGAATAATTGCTTCTTTAAAAGCAATACCGCTATGTTG
        TAATAAACGTATTATTGTAAAACTATCCACAACTTGCCATAAAATAACGAT
        AAGTTGACTCAATGCAAATATGGATATGGAAATAAACAACTGCTTCCATTG
20   AATGGAAGTATTATTAAAGCGATAGCATAACTTAAGTTTAAGTGGTTTTTT
        AAGTAATAAATATAACATTGAACCTAAAAAACCAATCGAAGATGCCAATAT
        AGCTAATGCTCCTGCTTGATAAATAGACCAGTGTTTCATTGAAAACATAAT
        AATTGCAACAATGATTAAACTAACTCTAATTACCTGTTCTATAACCTGGGA
        AATTGCTGGTATGGTCATTACTTGTTTTGATTGATAAAATCCTCTTAACAC
25   TCCTAAGACACCTATTAAAATAAAACTAAAACTGGCCATCTTTAACATGGG
        TGCTAAATTAGAATCGCCCATCCATCGGGTAATCATATTCGCAAACATAAA
        AAGCAAAATAAAGACGATAAAACCTATGCATTGTAATCGAAACATAACCCT
        TGTATAGACTTCATCGGATCGATTAACACCTATCACTTGAGTCACAGCACT
        TGGAATAGCATTCATAGATAAAATAACCCCTAGTGCTACGACAGGATATAT
30   TTGTTGATAAGCATATAAACCGTCATCACCTAAAACATTTTGATACGGAAT
        GCGATAAATGGCACTTAGTATCTTCACAATAATTAATGCTAAGGTAAGTAT
        AACGACGCCGTTAAACGCCGTCTCACTCTTAGTCTTCATCTTTAATCGCCA
        TACTTTCTTCAATACATCTAACTAAGAATTTCAAACTATCTAACCATTGTT
        TTGATTTTGTTAGCGTTACATTCATTGCATTATTTTGCACGCCAACTTTCA
35   TTGCTCTACCAAGAGGTTGCGTCTGTTTAAACAATTCTTCTCCATTAATAT
        CTTCAGTCGCTTTAGGTGATAAAATGATTTGTATAGATTTGCCTTTGTCTT
        TTATCAATTCGACACCCGCATGTAGAGCGTGGACTTTGATTTCAACAATAT
        CTAATAATCGTTCGACTTCAATTGGATAATCATTAAAACGATCTATTAATT
        CATCTTTGACATCGAAAAGTTGTTCTTCAGTTTCTACTTTTCGAAGTTTTT
40   TATAAATCTCAATTTTAGCCTGTTCACTTTGTATATATTCAGCTGGTAAAT
        AAGCATCTAAGTGCAATTCTACTTCAATATCTGGTGCATCCGGCGATTCTT
        CTTTAATGCCACGTTTTTCGTTTACTGCTTCTTCTAACATTTGAGAGTATA
        AATCGAAACCAACCGAATCAATAAAGCCATGTTGTTGCTTACCGAGTAAAT
        TGCCTGCACCACGAATATTTAAATCTCGCATAGCGATTTTAAAACCTGAAC
45   CTAGTTCGGTAAACTCCTTAATAGCTTGCAATCGCTCTTCAGCAGTCTCAT
        TTAACACTTTGTTAGCTGGATGTAAGAAATAAGCGTAACCAATTCTACTTG
        AACGTCCTACACGTCCTCTTAATTGGTATAGCTGGCTTAAACCAAAACGAT
        CAGCCTCTTCTATGATTAAAGTATTAGCATTTGGTACATCTACACCTGTTT
        CAATAATTGTAGTCGTTACTAAAATATCGTACTCGTGATTAATAAAGCTTA
50   ACATTGTTTCCTCTAAATCACGTTCAGTCATTTGGCCATGTGCTACAGCAA
        TGTTAGCGTCAGGCATTAACCTTTGAAGTTGTTCTCTTTTTTCATAAATGG
        ACTGCACTTTGTTATACAAATAAAATACTTGTCCATCGCGAGATAATTCAC
        GCTCTAATGCCTCTTTAATAAAGTTCGTATTCTGTTCTAAGACATAAGTTT
        GTACAGGAAAACGATTTTCAGGTGGTGTTTCAATCACTGATAAGTCACGTA
55   CACCTAACATACTCATATGCAATGTTCTTGGTATTGGTGTTGCAGTAAGCG
        TCAGTACATCAACGTTTTTTTTCAAAGTTTTAATGCGTTCTTTATGTCGCA
        CTCCAAAACGTTGTTCTTCATCAACAATAAGCAATCCCAAATCTTTATATT
        GAATATCTTTACCTAATAATTTATGTGTACCTACGACAATGTCAACATATC
        CTGATTTGAGCCCTTCTTTAGTTTCCCTTATTTCTTTAGCTGTGCGGAATC

```
GACTTACCAATTGAATTTCGACCGGAAAATCTTGCATACGTTCAAGTAGTG
TTTCGTAGTGTTGTTGTGCAAGAATCGTTGTTGGTACTAAAAACGCAACTT
GCTTACCATCCATGACAGCTTTAAACGCTGCACGGACAGCCACTTCAGTTT
TACCATATCCCACATCACCACAAAGAAGACGGTCCATAGGTCTTGCTCTTT
CCATGTCACCTTTAATTTCATCGAATGACTTACTTTGATCGGGTGTTAATT
CATATGGAAAATCATGCTCAAATGCTGATTGCTCTGCAGTATCTTGGCCGT
ATTGATAACCTACAGACATTTCACGTTCTTTATACAAATCAATTAATTCAT
CAGCTATATCTTCAACACTTTGTTGAACTTTCGCCTTTGTTTTTTTCCATT
CTGTCCCGCCTAATTTATTTAATCTAGGCGACTTATCTTCAGAAGCCACAT
ACTTTTGAACTTGATCCATTTGATCAACTGGAACGAATAATTGATCAGTTC
CTTCCACCTCGAACCTATTAACCTCGTCATCTTCGAGGGATCTTATAACCG
AAGTTGGGAAATCTCATCTTGAGGGGGGCTTCATGCTTAGATGCTTTCAGC
ACTTATCCCGTCCATACATAGCTACCCAGCTATGCCGTTGGCACGACAACT
GGTACACCAGAGGTATGTCCATCCCGGTCCTCTCGTACTAAGGACAGCTCC
TCTCAAATTTCCTACGCCCACGACGGATAGGGACCGAACTGTCTCACGACG
TTCTGAACCCAGCTCGCGTACCGCTTTAATGGGCGAACAGCCCAACCCTTG
GGACCGACTACAGCCCCAGGATGCGATGAGCCGACATCGAGGTGCCAAACC
TCCCCGTCGATGTGAACTCTTGGGGGAGATAAGCCTGTTATCCCCGGGGTA
GCTTTTATCCGTTGAGCGATGGCCCTTCCATGCGGAACCACCGGATCACTA
AGTCCGTCTTTCGACCCTGCTCGACTTGTAGGTCTCGCAGTCAAGCTCCCT
TATGCCTTTACACTCTATGAATGATTTCCAACCATTCTGAGGGAACCTTTG
AGCGCCTCCGTTACCTTTTAGGAGGCGACCGCCCCAGTCAAACTGCCCGCC
TGACACTGTCTCCCACCACGATAAGTGGTGCGGGTTAG
```


```
Sequence 3405
step.1002f05.cons.ok
GTGGTATTAAACCAGATACTATCAAAGATATTGTTGCTGAAGATCCAGATT
TAGTTATTGTTGGTGGCGGTATTGCGAATGCTGACGATCCTGTAGAAGCAG
CAAAACAATGTAGAGCAGCTATTGAAGGTAAATAAGATGAGTGAATTTAAT
AATTATCGTCTTATTCTTGAAGAGTTAGATTCTACTTTATCTCAAGTAGAT
AATACAGAGTATGAACGTTTTGCTAATGATGTTATAGGTGCAGATCGCATA
TTTACAGCTGGTAAAGGTCGTTCAGGTTTTGTTGCTAATAGTTTTGCAATG
CGCTTAAATCAATTAGGTAAAAATGCCTACGTTGTAGGTGAGTCAACAACA
CCTTCAATTAAAGAACATGATTTGTTTATTATTATTTCAGGTTCAGGTTCT
ACAGAACATTTAAGATTATTAGCTGAAAAAGCACAATCTGAGGGTGCAAAA
ATTGTCTTATTAACTACAAATGCGGAATCGCCAATCGGTAATCTTGCAGAG
ACGGTTGTTGAATTGCCTGCAGGTACTAAACATGATGTTGAGGGTTCGAAA
CAACCACTTGGTAGTTTATTTGAACAGGCTTCACTTATATTCTTAGATAGT
GTTGTATTACCTTTAATGGATGCATTTCACATTAGTGAAAAAACAATGCAA
GAGAATCATGCTAATTTAGAATAACTAGAATGAGAGATGAGCACTTTTGTC
TATGATTAGACAAAGGTGCTTTTTTTAATAAATGGTGCAACATTACTCCTTT
AAGTGTATATTTATAATATAACTTTAACTATGAGGTGAACATCATGAATTT
TGATAGTTATATTTTTGATTTTGATGGAACGCTAATTGATACAACAACATG
TCACGTCAAAGCTACGCAAAGCGCTTTTAAAAGATTAAATTTAGATGAACC
TACAGAACAAGCTATTTTACATACATATCATTTAAATTTATATAACAATTT
TAAAGCGCTAGCTTCACATGAACTGTCTTTTTATCAAATAGAAAAATTAAT
AGATGAATACAATCATTGTTTTAGCAACGATGAAATACATCAATCAAAAGA
ATATACCGGAATAAGTGAAGCATTAAAATTTTTACATAACCAAAAGAAAAA
AATATTTGTAGTGTCTAATAAAGAAATACTAACAACTCAAAAGTATTTAGA
TTATCTCGGATTAAGCCGTTTTATAACTGATTCATTAGGTGTCTGTATTAA
AAATGAAGACAAACTTCTTTGTGAAACGATTCAAAATTTGATACAGAAACA
TCATTTAATGATAGGTAAAACCGTGTATATAGGGGACACAGCACAGAATAT
CAAGAGTGCGAATCAAGCTCATGTGCAAACATGGCGCTGTCACATGGGGAGC
ACAATCTGCACACGAATTGTTGCATGAAAATCCTCATTATATTGTTAATGA
TCCAGAAGAATTTTTTAACAATTTTATAATATTTTTCAATGGTAGGATGATAA
TTCATGAGGAATTTTATCCTACTCTTTTTTTAAAACTATGAAGCAAGTATA
ATGTAAACATAATTATATAAAGTGTTTACAAATGGTGATGAAGATGAAAAA
AATAATGGAATATTTACAGCATTATATTAATCAATACCCGCATCGATTAGC
```

```
      TTTAGTGTTTGAAGATCGACATCTAACGTATGGAGAATTAAGTAAAGAAAT
      TTATCAGGCTAGTATGCGCTATAAAGAAGTAAAATTAAACAAAAAAGTAGG
      TCTAATGGATGAACATCCTGTAAATAATATTATTAACTATTTTGCGGTACA
      TCAAAGAGGTGGAATTCCTTGCATTTTTAATCATCAATGGAGTAATGAAAG
  5   GATACATCAACTTGTAAAAAGTTATGACATACAATGGTTAATTAAAGATAA
      TCATCTTACCTCAAATCATGATAACTCAATTTATAATGATGAGGTTATCCC
      ACGTAATGTTATACATATAGGTTTCACGTCAGGAACTACAGGTTTACCCAA
      AGCGTTTTATAGAAATGAACATTCTTGGATAGTTTCTTTTAAGGAAAATGA
      GAAATTACTCCAGCATTGTGAAGAAACCATTGTAGCACCGGGTCCTTTATC
 10   ACATTCACTTTCATTGTACGCATGTATTTATGCATTAAGTACTGGAAAAAC
      ATTTATAGGTCAAAAAAATTTTAATCCACTATCTCTTATGCGTCTTATTAA
      TCAATTGAACAAAACGACAGCAATATTTGTAGTGCCAACGATGGTACAACA
      ACTTATTTCAACTCAACGACATTGTTCATCGATTAAAAGTATTTTGAGTAG
      TGGTGCTAAACTTACATTGCAACAGTTTCAACAAATCAGAAATTTATATCC
 15   ACAAGCAAATTTAATAGAATTTTTTGGGACATCTGAAGCAAGTTTTATAAG
      CTACAATTTTAACCAATCATCTCCTGCTAATTCTGTTGGTAAACTTTTCCC
      TCATGTCGAGACACGATTATTAAATCAAGATGATGATGCAGTAGGATTATT
      AGCCGTTAGAAGTGAAATGGTGTTTAGTGGTTATGTTGGACAAAGCAATCA
      AGAGGGGGCATGGATTAAAACAGGCGACTTCGCTTATATTAAAAATCAACA
 20   TTTGTTTTTAGTAGGTAGAGAGAGTGATCGTATTATAGTTGGGGGGATTAA
      TGTATATCCAACAGCTATTGAAAGCTTAATTATGGATATTGAAGGCATTGA
      TGAGGCCCTTGTCATTGGTATACCACATGCTAAATTTGGAGAAATAGCGAT
      ATTGCTTTATTCAGGTAAAGTACAATTGAATTACCGACAAATTAAATCTTT
      TTTAATGAAACATCTTTCAAGACAAGAAGTTCCATCAAAATTAAAGAAAAT
 25   TGACCATATGATTTATACAGAATCAGGAAAGATTGCTAGAAAAGAGATGAA
      AAATAAATTTATTAATGGAGAGTTATAAAAATGAAACAACCTGTTATTATT
      GCAGCAAAACGTATAGCTTTCGGTAAGTATGGTGGCCGATTGAGGCATTTA
      GAACCTGAATCATTACTAGAACCTTTATTTAATCATTTTACAGATCAGTAT
      CCAAAAGTAATGTCTCTTTTGGATGACGTCATTTTAGGTAATACGGTAGGT
 30   AATGGGGGGAATTTAGCTAGAAAATCATTACTTGAAGCGGGATTAGATTTT
      AAAATACCTGGTATAACAATTGATCGTCAATGTGGCTCAGGTCTTGAAGCC
      GTTATA


 35   Sequence 3406
      step.1002f06.cons.ok
      TGAAGTTTAAGTCTTTTAATAAATCTACAACAAAAACTTTACGTTCTCCTT
      TGAAAAATTTACCGATTGCTGGGATAAATTTTTCATGATCTTTTAGCATAT
      ATATCAGGAAGAAAGGCACCATTATTAATAAAAATACAGTTGAGATGAAAG
 40   ACGTAATATATGACACTGAATTAGATAAAATAGACGTTGCGCCATCTCCCA
      TGGATTTAACCGCTTTATTAATACGATGTGTGACATCATCTGGTAATTTAT
      CCATTTGTCTTAACGAGAAATTAATTAATTGTTCGGCTTCTTTTTGTAATG
      ATGGTGTTTGTTTAATTAAGTTATTAATATTGGAAATAATGATAGGTGCTA
      TAAATGATACCACAATAGCGATGATAGCTATTAATCCTATGAATATTGTTG
 45   TTATACTAGCCCAACGTGGAAAGCCCCACTTTTCTAATATGTTTTGAAAAG
      GTAAACATATGTAAAAGAGAAATCCACTAATTAAAAATGGAAGGAAGACTG
      AACCAATGATAGTAGCTATTGGAGCAAATACTTCATGCACTTCCATAAATA
      GTTTGATGAGTATGAACAGCATAATAAAAAAATATTCCTGTTCTAAACCAAA
      CCTTATTGAACATAGATTGTGATCCTCCTATTTTATTCTTTTTCATACTCA
 50   ATAAGTATATATCAAAACTCAGTCATTAAAAAGAGTAAATACTTTTACAAT
      ATAACTGCATGCACTTTATTTTTAATAAAATGAAAAAACTAGTATGAAATT
      AATTTGTATACGAGCTTGCTATAAATTTAAAAGTATTATAAGTATTGATAC
      AATAGTATTAACATTACAAGGGAGCATCGGTTTATGTCGGAGTTTAAAGTA
      GGTAAGATTAATAAAAAAATAATACATAGTGATATTTTAAATAGAGATGTA
 55   ACATTATCGGTTTATTTACCTGAGGATTATACAAACTTATTTAAATATCAG
      TTGATTTTTTGCTTTGATGGTTTAGATTTTTTTAGATTTGGTAGAATACAA
      CGCATATATGAACAGTTACGTGAAGAACAATCAATACAACGTGCAATTATA
      GTAGGATTCCATTACGAAGATGTTGAAAAACGTAGGGAAGAATTTCATCCT
      TCAGGAAGTCGTTCTAATTTAACCATTAAAGCAATGGGAAAAGAAATTCTT
```

CCTTATATCGATGCGACATTTCCAACTTATAAAGTAGGTAATACAAGGTTA
CTTATTGGAGATAGTTTAGCAGGAAGTATCGCTTTAATGACTGCAATGACT
TACCCAACTATTTTTAGTCGAGTTGCGTTATTGAGCCCAATGTATAATGAA
AATATTAAGAAAAAAATTGATACATGTATGAATAAAGGTCAATTGACGATA

5   TGGCATGCCATTGGTTTAGAAGAAGCAGATTTTATTTTACCAACTAATGGT
AAAAGAGCTAACTTTTTAACACCTAACCGTGAATTAAATCAACTGATTAAA
GAAGATAATATTGAATATTTCTATAAAGAATTTAACGGTGGACATCATTGG
AAATCATGGAAACCATTGCTAGGAGATATTCTCTTACAATTTTTAGGTGAT
CCAATAAATGGAAAATATGTTTAATAAAGTAAGATAAATAAAATAGTATGT

10   TAGCGAATGGTTGTTGAATGCGAAACACTATTAATTTCAGAAATGTAATTG
TTTTCATGATAAAAGTAACGGTTTTCAAAAAGTTTTGATATAATAGGAATA
AGTTAAACAAAGGAGGAATTTAAATGATTTTAGGATTAGCATTGGTTCCGT
CAAAGTCATTTCAAGATGAGGTGAATGCTTATCGCAAGCGATATGACAATC
ATTATGCTCAAATAATGCCTCATATCACGATTAAACCTCAATTTGAAATCG

15   ATGATCATGATTTTAATTTAATTAAAAATGAAGTGAAAAATCGAATTTCTA
GTATTAAACCAGTAGAAGTACATGCTACAAAGGCATCTAATTTCGCTCCAA
TCAGTAATGTTATATACTTCAAAGTTGCTAAAACAGAGTCATTAGATCAAT
TATTTAATCAATTTAATACAGAAGATTTTTACGGTACAGCTGAACATCCTT
TTGTACCACATTTTACAATTGCCCAAGGTCTAACAAGTCAAGAATTTGAAG

20   ATATATATGGTCAAGTAAAATTAGCAGGGGTAGACCATAGAGAAATAATTG
AAGAACTATCGTTACTTCAATATAGTGAAGAAGAGGACAAATGGACTATTA
TTGAAACTTTTACATTAGGATAAAAAGTTCAAAATTGTAAAGTGAAATTGA
ATTTACAGAATCATTATTGTTAAATACGTGAAGAGCGCTTAATCAACTAAA
AACGCCAAATCCTATTGTGTTTCAGTGGGATTTGGCGTTTTTATATGAAAC

25   ATAATTATATTGTATAGTTAATCATTCAAAAAGATTAATGTACGTTTTATA
TTGAATATTCATATTTCAAATGCATATGAATGACGTTTCATTAATATATAC
AACTAATCTTTATATAAACGATTTATAGATTTTAAGGTTTAACCCTTCTGT
TTGTTTTAATAAAGTAATATCCGTAAAATACTGCAAGTGCAAGAAAAATCA
TCGCTGAAAAGTAAAAGGTATTATTTAAATTATTAGTAAATTGAGTAATTA

30   GACCTCCGACTAGTGGGCCTATCATTGAACCGAAGCCTTGAACACTGTTGA
ACACGCCCCATGTTTCTTCCTGTTCGTTAGGATTAATATGCCCAGCCATAA
AGGTATTCCAAGCCGGTAAGAGGATACCGTACATTAGCCCAATAAAAAGTC
CTATGGCCCAAACTATATATATATTTGTAATTGTAGATAGCCCGAATATAA
GAATTGTATATAGTATAAAGCCACTAAAAATAACTCCATACATAAACCCTT

35   TGCTATTATTGTCGATGATTTTTGATAAAAATAACATAGAGAAAGCACAGC
CTATGCCACCAATAATGATTGCTACTGTATATTCAACGGTTGATACTTTCA
CAACTTGCGTTGCATATTTTGGAAGAATAGGTACAAGTGCTGCTATAGCTG
CTCCTTGTAACAAGATACCCGGAAATAGAATAAGATGACG

40

Sequence 3407
step.1002f08.cons.ok
TAACGTGTGTCTCGCTATCAGTAATTACTGCAAAAGGTGTCACTTCAGAAC
CAGTTCCTGATGTCGTTGGTATACATATAAATTTTGCGTTTTTAGGTTTGG

45   TAATTTTATAAGTACGTTTACGAATATCTAAGAACTTTTGTTTTGCCCCAA
AAAATGAAGTTTCTGGATGCTCAAAGAACATCCATATTGCTTTGGCTGCAT
CCATTGCCGAACCGCCACCGAGTGCAATAATAGTATTAGGTTGGAAATTTA
TAAACATTTCTAACCCCTTATAGACTGTATGAGTTGATGGATTAGGTTCAA
CTTCGTTAAACACTTTGATTTGTGGTTGGTTTTCTCGGCGTCTCAGCACTT

50   GTTCAACTATATCAGTATAACCAATATTAACCATTCCTGGATCACAAACTA
TCATTACACGTTCAACATTATCCATCTCTGTCAAATACATAACTGAATTTT
CTTCAAAATAAACTTTAGGTGGGAGTTTAAACCATTGCATATTATTACGAC
GTTTTGCTATTGTTTTAATATTTAATAAGTCTACTGCGCTTACATTATGAG
AAATAGAATTTCTACCATATGAACCACAACCTAACGTGAGTGAAGGAATGA

55   GTTCATTATACATATTTCCAATTCCTCCGACAGCAGAAGGTGTATTTACCA
ATACACGGCAAGCTTTCATTTTTAGACCGAATTTTTGTTGTAATTGACTAT
CCTCGGTGTGAATTACAGCAGTGTGACCTAAACCACCAAATTTTAATATGT
CTTCACAAATTTGTAGTGCATGTCCTGTTGATTTTGCAGTTACCATTGCGA
GTACAGGTGATAATTTTTCACGTGATAAAGGATAATCTTTTCCAATTCCAT

```
      CAATTTCTGCGACTAATAATTTTGTTTTTTCTGGAACACTAATTCCTGACA
      ATTTCGCTATATCTACAGCAGATTTACCAACTATATCAGGTTTAACTGCAG
      TTTTATCTTCATTCATGATGGCATCTTCTAATTGTTGTAGTTCATTTTTAT
      TAACAAAATATGTTTGGTGTAATTTGAATTCTTTAACGACGTCAGTGTATA
  5   CTTCTTTATCAACAACCATGACTTGTTCAGAAGCACAAATCATACCATTAT
      CAAAAGTTTTAGAACCAATGATATCATTAACAGCACGTTTGATATGAGCAG
      TTTTTTCAATATAAGTAGGAACATTACCTGGACCGACTCCTAATGCAGGTT
      TACCTGTCGAATATGCGGACTTTACCATTCCAGAGCCTCCAGTCGCTAGAA
      CTAAAGCAATATCTTTATGATTCATTAATTGTTTAGTTGCCTCAATTGATG
 10   GCACTTCTATCCATTGAATACAATCTTTAGGAGCACCTGCTTTTGTTGCAG
      CTTCTAAAATGACTTTAGCAGCATATTTTGATGATTGTTGTGCACTTGGAT
      GAAATGCGAAAATAATTGGATTACCTGTTTTAATAGCAATCATTGCTTTGA
      AAATAGTTGTAGATGTAGGGTTGGTCGTTGGTGTCACTCCACAAATCACGC
      CTATAGGTTCAGCTACATACGTTAATCCTTTTTGTTTATCTTCACCTATGA
 15   TACCTACTGTTTTATTGTCTTTGATTGAATTCCATATGTACTCTGAGGCAT
      ATAAATTTTTGATAGCTTTGTCTTCATAAATACCTCTACCTGTTTCGTCGT
      AAGCTAGTTTAGCTAAATGCATATGCTGATCAACAGCAGCCATGCTCATCT
      GATGTACAATGTCATTAATCTCATGTTGTGATTTTTTAGATAGTTCTTTTA
      GAGCTTCTTGTCCTTTTTCTGCTAATGAATCAATCATTTGTGTGACTTCAT
 20   CTTTGTTTGATTCATATGTATTTTTTTTAGTTACAGATAACATATACAACC
      ACTCCTTAGATATTGTGAAATAATTCACAAACATTATAGTACAGCTCTTTC
      CAAATTAAAAGAAAATACAGACTATTAAAGAGTCAATTGTGATGTTTATAT
      GAACACATTTTTAACCTATATATTTAGTCGGGGTTGAGGTGTAAGACTTGA
      TGCAATGGGAACAAGACGTATTTGGTAGTTAAAAAAATCTATTTTTAAAAG
 25   AAGTAAATTTTAAAAAAATTATTGTAAGTATTCTTACTAAAAATAATGAGG
      GGTGTTGTAATGAAATTATCGTCTCACCCGACTAACGATGTTATCAGTTTA
      TAAAATAAAAAACTGAGACAATCATTTATGTCTCAGTAGCTCATGCTTTAA
      AAATAAATTGATAAAGAATAAGTGAAATTTATATTAATTATTTTGTTCAGT
      TGCAGTAAGAAGTTGTATCGCTTTGGGTAATACTTGAATCTCAATTGGTGT
 30   TTCTAAATTAATTTCGCCATCAATATCCACTTTCATACTAGGGTTTGTTGA
      GAGTGTGATGTGCTTACCTGATATGTGATCAATACCTTGTGTGATTTCGTT
      CCAATTCATACTATCACGTTTTTTTAATATATCATTCAATATATTTAGTGT
      TTGATCATTAAATACAAATGTGTTTGCTCTTCCATCTTGTGGCGATAAATC
      GGTTAGCGGAATTTGTCCACCACCTATATTGGGACCGTTTGCTATTAACAT
 35   CATCGAAGTATTGCCTTCTTTTGTTTCACCATCAACAGTCAATGAGAAATC
      AAATTTAACAGGATTTAATAACGTTTTAACGGTTGATCCAATATAGCTGAA
      TTTACCGAATATATCTTTAGAACCATCTTGAACATTCTCTGCATTTTGTAC
      TATTAAGCCAAGTCCAACGAAATTAAGTACATATAAGTCGTTCACTTTTAA
      CACATCATATGATTCAGCATGTGATGTTAATAATTGCTCACTAGCTGTTTT
 40   AAAATTAGGGTGCAGTTGAAGTGTTTTTGTAAAATCGTTAAAGGTACCACC
      TGGTATTACACCGATTGGTAAATTTAACTGATACTGCATAACGCCATTTAC
      TAGTTCATTAAGTGTACCATCTCCACCTAAAATAAATAAAACGTCTACATC
      AGAGCTATAATTTTCATTTTTAATAGATTTACAATATTTAATAATATCGCC
      TTTATTTTCACTGAGTTGAAGAGAGAGGTGTTTACACATTGAACTAAGAGA
 45   TTTTGCAACTTCTCCTATGCCATTATGTATATCTTTCAAACCGCTATGTTC
      ATTCAAACTTGGCACGGTACACCATTAAAACGTTTAGCGAATGATATGAAA
      GTGGTACGTATGCCGGGTACTACGACACCAAAGTATAAGCGTAATTTTAAT
      CGTGAAACATCACGTTGGGATTATTTAATTTCGCCAAATAGATATTCAACT
      GAAATATTTAGAAGTGCTTTTTGGATGGATGAAGAAAGAATATTAGAGATA
 50   GGTTATCCAAGAAATGATGTATTAGTTAATAGAGCCAATGATCAAGAGTAT
      TTAGATGAAATTAGAACTCACTTAAATTTACCTAGTGATAAAAAGGTTATT
      ATGTATGCTCCGACATGGAGAGACGATGAATTTGTGAGTAAAGGAAAATAT
      TTGTTTGAATTAAAAATTGATTTAGACAACCTTTATAAAGAACTCGGAGAT
      GATTATGTGATTTTATTACGCATGCATTATCTCATTTCTAACGCACTTGAT
 55   TTATCTGGTTATGAAAATTTTGCAATTGATGTTTCAAACTATAATGACGTC
      TCTGAATTATTTTTAATAAG
```

Sequence 3408

step.1002f10.cons.ok
CGCCACTGTGCTCTAAAGATATACGTGTCGCGCAACTTGTTTCTGTCGTGC
TTATTATGATTAGTGTGATATTCGTAATTTATAGAAGAGTTAAATACCAAC
CAATAAAATATGAAAATTCAGGACCATTAACATGGCCGATTAAAAAGGCTA
AGTGATTATTTTGAGAAGATTAAGTAAAGAGAAGCATCATCAAGTGAATCC
TTTATGGCATATATATAGAAAAATTAAATTCGTAAAAGTGCTAAAGCAAAC
GTTGATTATAGAAATTGCAAGATTTGTTCCTAGTATGAAACTTAAAAATAA
AATATATAAAAAGCTTTTAAAAATGGATGTCGGGAATCATACTTCATTTGC
ATATAAAGTGTTGCCTGATTTGTTTTATCCAGAATACATTTCGGTTGGCAA
GAATACAGTCATTGGTTATAATACAACAATATTAACTCACGAAGTACTTGT
TGATGAGTGGAGAGTAGGAAAAGTCATTATAGGCGATTACACTTTAATAGG
TGCAAATACGACAATATTACCAGGAATAACCATAGGAAATCATGTTAAAAT
TGGTGCGGGTACGGTTGTGTCTAAAGATGTTCCCGATTACAGTTTTGCATT
TGGTAATCCTATGCAAATACAATTAGATTCAGGAGGTGACAATGAATGGCA
CAAAAAGAAAATAACATCATTCCAATGATTTTTGATGAGGCTTTTTATCGT
AAAATGGCTACCCAAAAGTGGCGACAACAGGACTACAAGAAGGCGGCAGAA
TATTATGAAAAAGTGCTTGAGTTGTCACCTAAGGATTTTGATATTCAACAA
CACTATGCTCAATGTTTAGTTAAATTAAATATCGGCAAAAAAGCAGAACGT
CTATTTTATGAGAATATTGTTAAAGACTTTCATGTAGCAGAAAGTTTCTAT
GAACTAAGTCAATTAAATATAGAATTAAACGAGCCAAACAAGGCATTCTTG
TTTGGTATTAATTATGTAATATTGTCAGAGGATAAAGATTTTAGAGAAATG
CTTGAAAAAACATTCGAAGTCACGTATACAAATGAACAAAAAATTGAATTA
GAAGCACAATTGTTTTCAACACAACTTTTATTTCAATTTCTCTTTTCGCAA
GGTAGGTTAGAAGAAGCCCGAACATATATTTTGAATCAATCTTACGAGATA
CAACAGCATAGGGTGATTAGGAATTTACTTGCAATGTGTTATTTGTATCTA
GGTGAGTATGATAGCGCCAAAGCAATGTTTGAAGAACTTTTAAAGGAAGAT
AATTCAGACGTGCATGCACTTTGTCACTACACATTATTACTTTATAATAAA
AAAGAAACAGAAAAATATCAAAAATATCTTAAAATACTTAATAAAGTAGTA
CCACTAAATGACGACGAAACCTTTAAATTAGGAATCGTATTGAGTTATTTA
AAACAGTATCGTGCTTCTCAAAATTTACTTTATCCACTTTATAAAAAAGGT
AAATTTGTCTCTATTCAAATGTATAATGCATTGAGTTTCAATTTTTATTAC
CTAGGAAATAAAGACGAAAGTATTGAGATGTGGAACAAGCTCACTCAAATT
TCTGAAGTTGATGTTGGTTATGCACCTTGGGTAATTGAGGAAAGTAAAACG
GTATTTGAATCACGAGTGTTACCATTATTACTAGATGATAATAATCATTAT
CGACTTTACGGTATTTTTTTACTTCATCAATTAAATGGAAAAGAAATACTA
ATGACTGAAGATATTTGGTCAATTCTTGAATCAATGAATGACTATGAGAAA
CTTTATCTCACATATTTGGTACAAGGACTCACACTCAATAAATTAGATTTT
ATACACAGAGGTATGCAAAGGTTGTATAATTTTAAGAAATTCAAATATAAC
ACGTCTTTATTTACAGATTGGATTAATCAAGCAGAAATGATTATAGCTGAA
AATGTAGATTTAGTAGATGTCGATAGATATGTAGCTGCATTTGTTTACCTA
TCGTATCGTCGTTCTAGCCAACCACTTACCAAGAGGCAATTGATGGACGAT
TTTAATGTTTCTAGATACAAACTGAATAAAGCAATTGAATTTATATTGAGC
ATATAAATTTATGAAATGAAAAGTATGATATATAATGCGCATAATGATTAA
TAAGTAGGAGGCATTTAAAATGACTGAAGTAGATTTTGATGTAGCAATAAT
CGGTGCAGGTCCTGCCGGTATGACAGCAGCAGTATATGCATCTCGTGCCAA
TTTAAAAACTGTCATGATTGAACGCGGTATGCCAGGCGGTCAAATGGCAAA
CACTGAAGAAGTAGAGAATTTTCCAGGATTTGAGATGATCACAGGTCCTGA
CTTATCTACTAAAATGTTTGAACATGCTAAAAAATTTGGTGCGGAATACCA
ATATGGCGATATTAAATCTGTTGAAGATAAAGGCGACTATAAAGTTATCAA
TTTAGGGAATATAGTTGTTGATGTAAATGTGTGTCACAGTATTGTTTAACA
GTATTTTCATAACGTAATAACAATGTGTCAATCATCGGTGTTGCAACGTTT
AATGCTCTTCCAATCGCTTGAATTATCAAAGTACGATAATAATCTTCACTC
GGCATGCGTGGTATATGTATGACTCCTTGTTCATCAGTATCAACATGTTTG
TATGGTACGGCAGAAAAATCAAAATATGCACCTTGATCGTCCGGATTAGAA
AACGGATCGATTAAAATTGCTGTATATCGCACATAAAGTAGATACTCTTGA
TGAATAGCTGGTAAATTTTTAAAGTTTTCAATATCTACTTCGCGCATGGTC
TCATAACGTATAGGGTAGTTTTCTTTCACCATAAACTTTAGAAGATTGACC
GAAGGTACCTTTAATTTTTTTAATATCATCATCATTTCTTGCCACATTAAT
CGCATTTCGTGTATTAAGGTCATTGTGATTGGACCCTCTGGAAATAGCTTA

TATACATATACTGGTACTTTCGTCCCTTCAAATACCGCCTTTAATGAAAAT
TGATTCATAAACAATGGTGGGTGTACATAAAGTGAACTATTATGTATCTCC
GCATGTAGTGGTGTGTCCATCGTTGTTAATTCTATATTCAAATAGTCAAAT
AAAGACTTAAGCTTACACAACGTCATAGACTGAGATTGAGTCGAACCTACG
5 AATAATTTTGATTTAACTCGTGTGGTTAGGACACAATGTGGTTGTGCTTTA
TCAAATATTCGGGTATCGCCTAGATAAGTGGAAAATGAAATCACTTCACCT
TCACATTGAACATCTGATAGTAATTGCTTAACAAGCATATGTGATCCTAAT
GTTGGTGAGACCAAGATGATTTGCTTAATACGCTTTAATGTGGACTTAGAT
AATTGCTGTAATATCGGTCGATACGCATCGGCAGTACATGCTAAAATCACC
10 ACGTCATAATATTCAGTAATATCTTTAACATCTTTAAAAAAATGTCTAACC
GTAAACTTACCTGAAAAACACTGATGTGCATCATTTTGAGTCATTACTGAA
AAAAAGCCGTCACGTTGATAAGCATCAAAGACTCTCTTAGATTTGGTTGAT
GCATGAACGCGACTCACCATATCAACTATATGTTCTCCATGTAAATGACAT
AGTCGAGCAAGTTGAATAGCGACCGGTCCACTACCAACCATTAATATTTTA
15 GACATGTTCAACACCAACTTTTGATGCCACTTTACAATAAAGTGCAATATC
AAAAATCTGCTTCGGTCGTGTTTGTAGTTCTTCACATCGCCACTGTGTTTC
GTCAGTTGCTTGAGATGGATAATTAAACAATGCTTTTATATCATCACCGTA
TCGCATAGCTACAACAACTTCATCATTCGTCAATGTGTAAAGTTCATTCAA
AATGTCATATTTGATAGGTATCGTCGAACTGAATATAATATGTGTCACAGA
20 TTGG

Sequence 3409
step.1002f12.cons.ok
25 AAGATGTAAAAAAACAAAATGATTATTTAGATAAAATTTTACCTGAAGCTT
ATGCACTTGTACGTGAGGGGTCAAAGAGAGTATTTAATATGATTCCTTATA
AAGTACAAGTAATGGGTGGTATTGCTATACATAAAGGTGATATTGCAGAAA
TGAGAACAGGTGAAGGGAAAACATTGACTGCAACCATGCCGACGTATTTGA
ATGCTTTAGCTGGTAGAGGTGTACATGTTATTACAGTCAATGAATATCTAT
30 CAAGTTCACAAAGTGAAGAAATGGCTGAACTATATAACTATCTTGGCTTAA
CTGTAGGTTTGAACTTAAATAGTAAGTCAACTGAAGAAAAACGTGAGGCTT
ACGCACAAGATATCACTTATAGTACGAATAATGAACTTGGGTTTGATTATC
TTAGAGATAATATGGTGAACTATGCTGAAGAGAGAGTAATGCGTCCTCTAC
ATTTTGCAATTATTGATGAGGTCGATTCCATATTGATCGACGAAGCAAGAA
35 CACCTTTAATTATTTCTGGTGAAGCGGAAAAATCTACTTCTTTATATACTC
AAGCAAATGTTTTTGCAAAAATGCTTAAAGCGGAAGATGATTATAATTATG
ATGAAAAAACCAAAGCTGTACATCTTACAGAACAAGGTGCAGATAAAGCTG
AACGTATGTTCAAAGTAGATAATCTTTATGATGTTCAAAATGTGGAAGTGA
TTAGTCATATTAATACAGCTTTAAGAGCTCATGTTACTTTGCAACGCGATG
40 TTGATTACATGGTCGTTGACGGTGAAGTATTAATTGTTGACCAATTTACTG
GACGTACAATGCCTGGACGTCGTTTTTCTGAAGGTTTACACCAAGCAATTG
AGGCTAAAGAAGGTGTAGCAATTCAAAATGAGTCTAAAACGATGGCATCCA
TTACTTTCCAAAACTATTTCAGAATGTATAATAAGTTAGCGGGGATGACTG
GTACAGCGAAAACCGAAGAGGAAGAATTTCGTAATATCTATAATATGACAG
45 TTACCCAAATTCCAACAAACAAACCTGTTCAACGTAAAGATAATTCAGACT
TAATTTATATTAGTCAAAAAGGAAAGTTTGATGCGGTAGTTGAAGATGTTG
TAGAAAAACATAAAAAAGGACAACCCGTCTTACTAGGTACTGTTGCTGTTG
AGACTTCTGAATATATTTCAAATTTACTAAAAAAACGTGGTGTCAGACATG
ACGTATTAAACGCTAAAAATCATGAACGCGAAGCTGAAATCGTTTCAAACG
50 CGGGGCAAAAAGGTGCAGTTACAATTGCCACAAATATGGCTGGACGTGGAA
CAGATATTAAACTTGGTGATGGTGTTGAAGAGTTAGGTGGACTTGCTGTTA
TTGGTACTGAGCGTCATGAATCAAGACGTATTGATGATCAATTACGTGGAC
GTTCAGGACGCCAAGGTGATAGAGGAGATAGTCGTTTTTACCTATCTTTAC
AAGATGAATTAATGGTACGTTTTGGTTCAGAACGCTTACAGAAAATGATGA
55 ACCGTTTAGGAATGGATGATTCAACGCCAATCGAGTCGAAAATGGTATCTC
GAGCTGTAGAATCAGCTCAAAAACGAGTAGAAGGTAATAACTTTGACGCGC
GTAAACGTATTCTAGAATACGATGAAGTTTTACGTAAGCAACGTGAAATTA
TTTATAATGAGCGTAATGAAATCATTGATAGTGAAGAAAGTTCTCAAGTCG
TTAACGCGATGTTACGTTCTACATTGCAACGTGCGATTAATCATTTTATTA

ATGAAGAAGACGATAATCCTGACTACACGCCATTTATCAATTACGTTAATG
ATGTGTTCTTGCTGAATTATTAATGTTAAATGCTATCCATGTGCCATTACA
CACCGGTAAAATTGAAGAAATGACACGTGTACTTCGTGAAAAATAAGCATA
ATATATTTAAGCGTCTGGATTATCATCCCCAGACGCTTTTAAAATGTTTAG
AGAATTATCAAATTTTATTTGAATGCTTTCGCTAAAACAAAGATAAACATA
TCTTTCGAGGAGTTTTATCTAGTATACGTACATTCTTCCCTCAACAAGCTA
AAAACATAACTATCTGAGAATTCCCCTTGTAATCGTTCATTACTTCTCAAT
ATACCTTTGCGTGAACCCAAGTCTAATTGGTATAGCTCTGCTTTTCTTATT
TTTAGTAGACATTCGTATTTCTATACGATTAATATCATACACTTCATATAC
ATAGCGAATTAACGCTTTAGTACATTTAGTCATAATACCTTTCTTTTGAAA
GTCTTCAGCTAAATAATAACCAATTGAAGTTGTTTTATTAACTAAATCTAA
GTAATGCAATCCTATGACTCCAATCAATTCTTTATTACTCCATATTCCACA
ATGAAATCCATTACCATCGATAAATTGTTGCAACGCCGAATGGATAAAGTG
TTTACTATCTTCAACTTTCTTCGTATGTTCAACAAAAGGCAGAAATTCAGC
TAAATAGTCACGATTGCTATCTACTAATTTAAATAACTGTTCGGCTTCTCG
TTCTTCTAATATTTTTAATTTAATATTTTTATTTATTTGATAACTAAACAT
ACAAAATCATCCTTTCTTAGAATTTTCTGTCATTTTAGTATACGCGAGTAT
AGGATTATGTCATAAAACTTGATTACTAGAATTCGATATTCTTTAAGTTAC
AAGTGTGTATGTAGTATTTCTCTTGCTTCTATTGTCTTATACTTACAACCA
ATATTAGAGTGAAAAAGCTTAACTTAAGATAGATAATCATCAACAACACAA
ACGTTAAAAACACTGAGAAAACTTTTTTAAAATATACTTCTCAATATTGAT
TAGTAAAATCCAAAATATGAATAATAAGTAGACATTAAACAATCTTTATTT
TATTCTATAAATAATATAAAATAAATTATTAGATTAATACATTATTTTTTT
ATTTCTAATTTAAGTTGAAATCGTATGTAATCGGAGGAAACCATAATGATA
TATTCTATCGGTAAGAATTTAGGTAATAAATTAACAGGTATAGAACAAGCT
ATGATCAATAGATTAAAGCTATTTAAAGATAATTTAGTCCCAAATAAACTC
ATATTCACATCTTGGTCACCACGTTTATATATGCATGCACATTCGTTAAAC
ATCGATTCAAAAGATATTTTCAGTCTTTACGATTTTCTACAAGATAGTATT
AACTTTGAGAAAAAACATATTGATTGGATAAATTATTGGCAAAATATATGT
AATTATACCTTAAAATTCGTTGAAAATACGAATGATATTAAAAATATACGAT
AACGACACATATAAAATGTATGTGCATTTTGTTGATTCAAATTATCAAACT
TTAGACTATATTAACCATTTTGATATACAACAACGTAAAATTCGAAGAGAT
TTTTACGATACAAGAGGCTTTTTTAAGTTGTAGTAGAATTTTAACCTCTCAA
C

Sequence 3410
step.1002g01.cons.ok
TAACCGTAGGTATAGTGGTTAGCTTTACATAATTCAATGACTTTTTGTGTT
ACGGTATCATTTTTATAGGTTTTGTTTTTCCATCGGTAATATGTTGATTTA
GGTATGTTTAATACTTCTAGTATCAATTTGATTGAATATTTTACTTTTAAT
TGATCCACTAAATCTATGACTACTGTTGGTACCACTTCCTTTCCAATGCCT
TGTACTTTTTTAAAATATCCAATTCTATATCTTTTCTCTTATTTTCTAATT
TTAATTGTTCTACTTCTGACAGCTCTTCTAATCCTTTACCGTAGGTATATT
GTTTACCAACTTGTTGTGAAAATCTATAACTTTCCCCATTTCGATACCATC
GCCACCAAGTTTTCACTTGCGTCCTATTTCTAATATTTAATTCTTTCATAA
TTTCTTTTGTTGAAAATCCTGCTGCTTTCATTTCAACTGCTTTATACTTTG
TTTCTACTGAATAAGAAGCTCTTTTCATAGAAAAAAACCTCCGTATGATTC
ATTTTAATATGAATTCAACGAAAGTGTTTTTATATAATTCCCACAAATTGG
GGTCAGTCTAATAATGAATTTCGTGATTGTAGGCTTTTTATTTTATATATT
CATTAATTTCTCTTCTTGACGGATACGTACAATTAAAAAAATAAGCATATGG
AACTAATAAGAGTGTTGTGTATGTAGCATTTGTTAGTAGTAATACACCAAT
TAATTCGGGAATGATATTTAAAAAAATAGTTTGGATGTTTCGTTACTTTATA
TAACCCTGACTTAATAATAGGGTGATTGGGCAGTATAAATAATTTCAATGT
CCAAATACGACCTAATGTCTTAATAACTATAAATAGCATGATATAAGCGAT
GATCAATATGATTAAGCCAATGCCGTTAAGTAGACTAAATGTATCCTTACG
AATGAATGCTTCTATAGCTGCACTCATGTAAATTAATACATGCGTAATGGC
TAGATATTTTGAATTTTTCACACCATATTCCACCGCGCCCTCTACCTTTAG
CTGTTTTGCGTGTTGCATAGATATCTTTAAGCTGATGAGTCGAATACAGAA

AAAGATAAATAAAATAGTTAAAATCATGATGTCCTCCATTGTTTGTTATAT
ATATATTCTTTTGCTCGTACAATACAAAGTATTTTATATAGAAAATGTTTC
AAATGCCATTAACTATTTATTAACTTTAGCACCAACTCATAACTCTTCTAA
CACCTCACATATTAAGTTCAGTTTTTCGGATAATTTAATACTTTTAAGGAT
5   ATTAAGCGCTTACATTGATGTGATATATTTTTTTTTAACGAAGATGATAGGG
GTGTTAGGGATGAACTTTTTTGATATTCATAAAATGCCAAACAAAGGGATA
CCATTAGCTGTACAACGCAAATTATGGCTCAGAAACTTTATGCAAGCGTTT
TTTGTCGTATTCTTTGTTTACATGGCGATGTATTTAATTCGAAACAATTTT
AAAGCGGCACAACCGTTATTAAAAGAAGAAATCGGATTAACAACATTAGAA
10 CTAGGTTATATAGGATTAGCGTTTAGTATTACTTACGGTTTAGGAAAAACA
ATACTCGGTTATTTCGTTGATGGGCGTAATACGAAACGTATTATTTCCTTC
TTATTAATATTATCTGCGATTACAGTACTTATTATGGGATTTGTATTAAGT
TATTTCGGTTCTGTGATGGGGCTATTAATTGTATTGTGGGGGCTTAACGGT
ATATTTCAATCTGTGGGTGGGCCTGCAAGTTACTCAACGATTTCAAGGTGG
15 GCGCCTCGAACAAAGCGCGGTCGTTATTTAGGCTTTTGGAATACATCACAT
AACATTGGTGGTGCTATTGCTGGTGGTGTCGCACTTTGGGGCGCGAATACA
TTTTTCCACGGTAATGTGGTTGGAATGTTTATTTTTCCTTCCGTCATCGCT
TTAATCATTGGGATTGTGACATTATTTATTGGTAAAGATGATCCAGAGGAA
TTAGGTTGGAATCGTGCCGAAGAAATTTGGGAAGAGCCTATCGACCAAGAA
20 AACATTGATTCTCAAGGTATGACTAAATGGGATATCTTTAAAAAATATATC
CTTGGAAATCCTGTGATTTGGATTTTGTGTATCTCTAATGTTTTTGTATAT
ATCGTGCGTATTGGTATTGATAACTGGGCACCGCTATACGTATCAGAGCAT
TTACATTTTAATAAAGGTGATGCGGTGAATACTATTTTTTACTTTGAAATA
GGTGCATTAGTAGCTAGTTTATTGTGGGGCTATATCTCAGATTTATTAAAA
25 GGTCGTCGTGCGATTGTAGCGATTGGATGTATGTTTATGATCACCTTTGTT
GTACTCTTTTATACCAATGCAACAAGCGTGACAATGGTCAATATTTCTCTA
TTTGCATTAGGCGCTTTAATCTTCGGTCCACAGTTACTCATTGGTGTATCT
CTGACTGGCTTTGTTCCTAAAAAATGCAATTAGTGTCGCTAACGGTATGACA
GGTTCATTTGCATATCTATTCGGGGATTCAATGGCTAAAGTGGGTCTGGCT
30 GCAATCGCTGATCCAACACGTAATGGTTTAAATATTTTTGGGTATACGTTG
AGTGGTTGGACAGATGTCTTTATTGTATTCTATGTAGCTTTATTCTTAGGA
ATGATATTATTAGCCATTGTTGCTTATTACGAAGAAAAGAAAATTAGAAAA
TTAAAAATTTAAGATTAAGTGAATTTAAATATATACTCCCTACTATAAATT
TTATCATTATGGTAGGGAGCTTTTTCGTGATTGATTTGCACTATTTTGATT
35 GCTTATTATAATCACTTGGTGACATATGTAAATATTTTTAAAATGATAGC
AAAACATTTTATACTCAGAAAAACCTACTTTTTCAGCAATTTCAT

Sequence 3411
40 step.1002g02.cons.ok
TTAAGTTGTTGTAAGTTATCTACTTGTTCAATCCATTCGCTATCTATCGCT
TTCAAAATTGATTTTTGAATGAAACGAAGTTTTAAATATGAATCAGCAGCA
ACTTCTAAACGATTGTTAAATTGTTGAGTAAATTGTTGTATTAAGAATTGT
ATGATTTCTTCATCATTTTGCATATTAATATTTGATACATCTTCATCGAAG
45 ACAAAACTTAAGTTTTCGTATATATAATTCACAAGTGCACGTTCACTACTT
AAGTCAAGATTTTTAACGTCTTTTGTAAACACATCTCGTGCAAGCTGTTCA
AAATTAAAATCATCAAAATCGCTTGCTTCAAGTATGTGATTACGTTCAGCA
TAAATTTTATCTCGTTGAACACTAATACTTTTTTCGAATTCATTTGCCATT
TCTCTATTTTTCATAGCAGTCTCTTCAGATACACGTTGCGCTTTATTAACA
50 ATTGACTTTACACGTTTTTTAAAGAGTGCACTACTTTCTAGTTTAGATGCA
TCCATCGTTTGGAGGTTTTTATTTTCTGCCAAGTTAGAGTTACTCCAACGT
TTTACTAAATCATCATCAAGTGATACAAAAATCTGTGAATATCCAGGATCT
CCTTGGCGACCTGAGCGTCCTCTTAATTGACGATCAACACGGCTATTATCC
ATATGTTCATTAATAATCACTGCTAAGCCACCGATATCATGAACCTCTTTT
55 GATAACTTTATATCAGTTCCACGCCCTGCCATACTTGTAGCAACAGTAACT
GCAGATAATTGTCCCGCCTCAGCAATCATTTGTGCCTCTTTAGCTACATTT
TGAGCGATTAATAAATTGTTGGGTATATCACGTTTAAATAACTCAGCTGAA
AAATATTCTGCCGCTTCTGCAGTACGTGTAATTAGTAACACAGGTTGTTGA
GTTTCATGTATACCAATCACTGTCTTTAAAAATTGCATCGTTCTTTTTGTCA

```
CCATTAGCAAATACTCTATCAGGTCTATCATCTCGTTCAATCGGACTGTGA
GT
```

5   Sequence 3412
       step.1002g03.cons.ok

```
GTATTCGGTAATATCATCTGTAAGATTTTTAATTTCTTGTCTTGCATTTTT
CATTGCAACAGTGTGATGGCCAAAATTAAACATAGCTCTATCATTATCACT
ATCACCAATAACTAAAGTTTCTTCTTGCTTAATACCATAGTGATCAATCAT
TTCCTTGATACCTGTACCCTTATCTGTATGATATGGCATCGTTTCAGCATT
ATAACGCGAAGAATTAGATACGGTAATACCTAAATTTGATTGGTTTTCAAC
AAGCTGGTTGCGAAAATCAGTTATTTTATCTAAGTTGGGACTAAATAAATA
TATTTTAGAAAAGTGTGCATCAGGTAAGGTATCTCGCCAATCTATTTTTCC
TTTAATTGCCTCTCTTCGCGATGACCACTCACTTTGACTTACAGCGTCAGG
TGGCTCTATAGTGGAAATCATTTCTTTCATCCAATCTTCATCTTCTTTAAG
AGATATACGATTACTTTCAAAAGGAAAAACTTCATAATAAATATGTTGTTT
TTTAGCCAATTCCACAATTTTCTGTACTCGTTCTAAAGTTAAACTATGTCT
AAACAAATTATCTCCATGAATTTCACCTGAAGTTCCATTCGAACTGATAAT
ACCATCTACAGTGAATCCATCAGGAACAAGCTGACTGATTTCTGAATAAGA
TCTCCCAGTAGCAAGGAAAACTTTATAATTTTGCTCTCTCAATTCATTAAT
CACTTGTTTAGTATATTCCGAGGCTTTATTATTTTCATGTAAAATCGTTCC
ATCCATATCTAAAAAAATTGCTTTTATATTTTTCATATTATCCTCCATTCA
AGACTCTAACTTATACACTTAAATTGTGTCACTTATCATACTAAAGTGCAA
CGGCTTTGATTACACTTTGAACATACTATAACGTTGTCTTCTCTTATTTAC
AATGACATCCCTATCAAAAAAAATTGTCATCCGATAAAGATCTCATAACATC
TTAAATATTAACCTTGTTAAATATTTAAAAGTCTTTAAGTAAATCTTTTTC
AAAAATAACGGCATTTCTTCTATTATATGTGTTACAAGTATGACAGAAAGT
TGTGAATCATTTAAAAATATGCTTACCAAATGTTTTAAATAGTTGTCTCTTT
ATATAAAATCCATTCCAGTTGTAGGTTTATCCAATACTAGTTAATTTAGGG
TACTCATTAAAGTCATTTAACAACGGCATTAAACCGCTTCGTTAAAATTAA
TTGAAGGATAAATCTAAAATTAATTCAAAAATAAAGAGTAGCCATTTAGGC
TACTCTAAATTATAAACATTTTACTCTACTGTAACCGATTTGGCTAAGTTA
CGAGGTTTGTCAACATCTAAATCTCGTTGTAACGCAGCATAATATGAGATT
AATTGCATTGTCACTACAGATACTAAAGGAGTTAATAATTCATGTACTTGT
GGAATCACGTATGTGTCTCCTTCTTTATTCAAACCTTCCATTGAAATCATA
CAAGGATATGCACCACGTGCTACTACTTCTTTCATATTTCCACGAATTGAT
AGATTAACGTTTTCTTGTGTAGCTAAACCTATAACAGGTGTGCCATCTTCA
ATCAAAGCGATTGTTCCGTGCTTTAATTCCCCACCTGCAAATCCTTCAGCT
TGAATATAAGAAATTTCTTTTAATTTTAATGCACCTTCTAAACTAACATTA
TAATCAATTGTTCGTCCAATGAAGAATGCATTACGAGTAGTTTTTAAGAAA
TCCGTTGCAATTTGCTCCATCTTAGGTGCATCGTCAACAATTGTTTCAATA
GCTGTAGTAACCTTAGCTAGTTCTCTTAATAAATCAACATCGGTTTCACGA
CCATGATTTTTAGCAACAATTTGAGATAAGATAGATAAAACAGCAATTTGC
GCTGTATATGCTTTTGTAGATGCGACTGCAATCTCAGGTCCAGCATGTAAA
AGTAATGTATGATCCGCTTCACGTGATAATGTTGAACCAGCAACATTAGTA
ATTGTTAATGATTTGTGACCTAACTTATTTGTTTCAACTAATACAGCACGA
CTATCAGCTGTTTCACCAGATTGTGAAATATAAATAAATAGTGGTTTTTCA
GAAAGAAGTGGCATATTATATACAAATTCAGAAGCTACATGAACCTCAGTA
GGTACACCTGCCCATTTTTCAATAAATTCTTTACCAACCAATCCAGCATGA
TAACTAGTACCAGCTGCAACGATGTAAATACGATCAGCATCTGCTACATCA
TTAATAATCTCTGAATCGATTTTTAAATTACCTTTTTCATCTTGATATTCT
TGAATAATGCGACGCATCACTGCAGGCTGTTCATGAATTTCTTTTAACATG
TAATGATCATATACGCCTTTTTCTGCATCAGCAGCATCTATTTCTGCCGTA
TACGTATCACGTTGTTGAATGTGCCCCTCAAGATCTTTAATTTCTACTGTG
TCTCGCTTAACAATAACTATTTCATGGTCATGTATCTCTTTGTATTGGTTT
GTAGTTTGTAACATTGCTAGAGCATCAGAAGCAATAACATTGAAACCTTCA
CCTACACCTACTAAAAGCGGAGACTTGTTTTTAGCCACATAAATAGTATCT
TTATCATTATCATCTAATAATCCTAAAGCATATGAACCATGTAATAATTTA
ACTACTTTTGTAAATGCATCTTCTGTAGCTAATCCTTGTCTAGAAAAATAA
```

```
TCTACTAATTGTACAATAACTTCCGTATCAGTTTCTGATGAAAAAGTGACA
TCAGATAAATATTCAGCTTTTAATTCTTCATAATTTTCAATTACACCATTA
TGAACTAATGTAAAACGTCCAGATGTTGACTGGTGTGGATGTGAATTTTCA
TAATTTGGAACACCATGTGTCGCCCAACGTGTATGACCAATTCCTAACGTA
CCATCCTCATCACTATTATCTGCAAC
```
5

```
Sequence 3413
step.1002g04.cons.ok
```
10
```
ACTGGTCCACACGTACACTTCCAAGTTGAAAGAGGTCGCCATGATGACATC
ACAAACAGAGGGACAGTAAACCCTGCTAAATGGCTCAAAGGTCACGGTGGT
GGAAAAGTTGGTGGTAGTGGTTCTGTAAACGCACGTAGAGCAATTCAAAGA
GCACAATCTATTTTAGGTGGACGTTATAAATCGTCTTATATTACCGAACAA
```
```
ATGATGAGAGTTGCCAAACGTGAGTCTAACTTCCAATCAGATGCGGTTAAT
```
15
```
AACTGGGACATCAACGCACAAAAAGGAACGCCTTCTAAAGGTATGTTCCAA
ATGATTGAACCATCTTTTAGAGCATATGCTAAACCAGGACACGGAAACATC
TTAAATCCAACTGACGAAGCTATATCTGCTATGCGTTACATTGTAGGTAAG
TGGGTTCCTATTATGGGGAGTTGGAGAAGTGCATTTAAACGTGCTGGAGAT
TATGCTTATGCTACAGGCGGGGTTATTAACACTGCTGGATTATATAATTTG
```
20
```
GCAGAAGATGGATACCCTGAGATAGTAATCCCTACAGATCCAAGCAGACAA
TCAGATGCGATGAAATTGTTACATCTTGCTGCAAGTAAAATTAGTGGAAAT
AACAGAAATAAACGACCTAACCAATTACGTACACCTAATGTTACTAGTAAT
ACAGTTGATAATGCAGAATTACTACTACAAATGATAGAAAATCAACAGAAA
CAAATAAACGTGTTAATGGAAATAGCACGAAGTAATAAAACTATTGAAAAA
```
25
```
CAACCGAAAGGTTTTTCAGAACGCGATGTAAGTCAGGCACAAGGTTCAAGG
TTAAGACTCGCTGCTTATAGCCAGGGAGGTTTATAAATTGGAAAATAAAAA
AGTAAAAATATTTAACGATCATTTCGAAGAAACACTAACGGATATTCCTCA
TCTTAAGTTTCACAGAGGATAAATATCACATACAGATTTTGCTATACGCAA
CATTGTTAAAACATGATAACTCAATCCACTCGCGAAATTATGATGATGAGA
```
30
```
ACTAGCTGCTGGATAAGTGAAAAAACGATCTTGATACTTTTTAATTAAATG
TCTAGTAATGCGTTGTAAGTTAGCATTTTCAATATCTAACATAAAATGCGA
TAATTCCTCTTGTATTTCTGTAGGTGACATTGGCGCACCGTCAACAAAGTC
TTTAGTTGACACTTTGTCTTCAGCTTGTGCAAGACGTATTTGATGTATTTT
CATCTGTTTGCGTCCACGATAATTGATGATATCACCTTTGACATGAACTAT
```
35
```
TGTTTCTGGTTTTAAAGTTTGCATATCATCTTTCGTAGCAGTCCATAATTT
AGCTTCAATATCACCACTTTTATCTTGCAGAAATAGTGTCATGTAATCTTT
ACCCTGAGCTGTAACACCTTGAGTTGCACGATGGATCAAGAAAAAATGATC
CACTGAGTCGCCGGGATTTAATTTTTCTACATTTCTCATTTTTTCCCGCCT
TCCTGTAATTTATTTAAAGTTAATACTTCTTTAGCCGGTATGACATGATCT
```
40
```
TTTGTACAAGTAAAATAAAGTATTTGATAATGTTCTGATAGTTCTCTCAAA
TATTTCAACATACGTTCTTTACGATATTTATCAAAATGAACAAATGCATCA
TCTACAATCACTGGGAATGGATAATACGGTTTTAATACTTTAATAAGACTA
ATACGTAAAGCCACATATAATAATTCTTTTGTAGATTGACTCAACTCAACT
GGCTCAAATACTTGTCCGTTAGAATGCTTTACATGTATTTTATGATTTTCA
```
45
```
GTATAATGAATCATATTGTAAGTACCATTTGTTAAATTTTTAAAAATAGAT
ACAGCTTCATTAATCACTTGTGGTAGACGCTTATCTTTTATTTGCTTGATA
TGTTCTTCCACTAAAGCTTGCATATAACTTAAGCTTGCCCAATCCTTAGCA
ATATCGTTAAGTCTATTTTTTAAGCTATAATATTCATGTCTTAATTGTGCT
AAAGTTCTATCTGTTTCCATATGATTAATCTGAGCATTTAAATCACTAACT
```
50
```
TCTGCTTGCATTTCTAAAAATTGATCATTGTAATCGTCAACTTGTTTAGCC
AATCTATGATCTTCTTCTTCTAGTTGAGCAGTAGTTTTCTCACTTAATTGC
GAACTCATTTCATAAGTGTAATTTTGATTTTCTAAATATTGATTTAAATCA
TTAAATCGATTTAAATCACTTTGATATGTTTCAAAATGATGATGATGTGTA
TAATAATCTTCTTCATTATCTACATCAACATAATCAAATAATTGCTTAATC
```
55
```
ATTTGATTCGTTTCACTTAAGCGTGACTTCAATTGCTTTAATTCATTATTG
AGTAATTGCGTTTCAGTTTGATTTTTATTCCAAGCCTCATTTTGTTCTTCT
GCATTTTTTAACCACTGTTTCACATCATGGAAGAAAGACATCTGATTAAAT
AGTTTTAAATGTGGTTCAGTTACAGATTGTGCACGTTCATAAAAGTGATTG
ATATCTTTTAGCAGATTATTGCGTTGTTGATTTAAATCAATAATGTATTTA
```

TCATGCGCTTTAATTTGACCAATGGTTGAGATACTTTCAACCACTAATTCA
TCAGAAAGTTTTTCTGATAAACGTAACTCTTTTTTTTATTTCAATAATGCTA
TGTTTTAAACTATCTAATCGCTCATTTGCAGTCGTTAATGATTGATTGATA
TATTGATGTTTTTCTTCAAGTATTTTTTTATTTTTTTTTAGCATGACGCCAT
TGTTCACGAACTTGTTGTTGATATTCTAAATCGAAGTCAAGATTATATTCT
TTTTCAAGTTGCGTGAGTTGGTTTTCTAAATCATTAATTTCCTGACTTATT
GCTGTGCTATAATCTACTGCTTTAGATCTAGAAAAAAATGATACCTACTACA
AAAATCACAGTTAATAGAGCAAATATTATACCAAAAATAAGATTTGCAGTG
AAAAAAGAAAAAATAGAAAGTGCTGCTGATAGAATAGTCAAAACAATAAAT
CCTATTCTCAAAAACTTTTGTCTTTTATTTTTTTGTGTTTGTTCTTCTTCA
AAAGTTTCTTTTAACTTTTCATACAAGTTCTCTTTTTCATGTAATTCTAAA
ACTTGTTGTGTATATTCCTTTTTCTTTTCAAAGGTTTCGTCAGGAACAAGC
TCATTCTCAACTTGATTAATCTCATTAGAATTAGAGTTTCTTTCA

Sequence 3414
step.1002g05.cons.ok
ACTTTGTTGATTCAACAGTCTCTGTAGCTAAAGAAGTTCAAAAGTCCGACG
ACAATCTAGGTATTGCTATAGATGCCTATGGTGCAGGTAGTTTTATTGTAG
CAACTAAAATTAAGGGAATGATTGCGGCTGAGGTTTCAGATGAGCGTTCAG
CTTACATGACACGTAGCCATAATAATGCACGCATGATTACTATGGGAGCTG
AAATTGTTGGAGATACGCTTGCTAAGAATGTCGCAAAGAATTTGTCAATG
GTCATTACGATGGTGGACGTCATCAAATCCGCGTAGATATGTTAAATAAAA
TGTGTTAATAGGAGGAATAAAAATGAAAATTGCAATAGGTTGCGATCATAT
TGTTACTGATACAAAAATGGAAGTTTCACAACACTTAAAATCACAGGGACA
TGAAGTGATAGATGTTGGAACTTATGATTTCACACGTACACATTATCCGAT
TTATGGAAAAAAGGTAGGAGAAAAAGTTGCGAGTGGTGAAGCAGATTTAGG
TGTATGTATTTGTGGTACTGGTGTAGGAATTAGTAATGCTGCAAACAAAGT
ACCAGGTGTTAGAACTGCTTTAGTTAGAGATATGACATCAGCGCTTTATTC
TAAAGAAGAGTTAAACGCCAATGTTGTAAGTTTTGGCGGTAAAGTAGCAGG
TGAATTATTTATTTTCGACATCGTTGATGCATTCATTGAGGCAGAGTACAA
ACCTACTGAAGAAAATAAAAAATTAATTGCTAAAATCAATCATTTAGAAGC
ACATAACAATGACCAAGCTGATCCACATTTCTTCGACGAGTTCTTAGAAAA
ATGGAATAAAGGTGAATATCACGATTAAGGAAGATGTCCAATGATTTTAAC
ATTAACTTTAAATCCCTCAGTTGATATATCTTATCCTTTAGATCAGTTTAA
TTTAGATACTGTTAATAGGGTATCTCAAACAAGCAAAACAGCAGGCGGTAA
AGGATTAAATGTTACTAGAGTATTGTCTGAATTTGGTGAAGACGTAATAGC
AAGTGGTTTTTTAGGTGGAGCATTAGGTCAATATATTGAAGAACAAATAGA
AACTACTCGTATTAAACAAGCATTTTTCAAAATCAAAGGCGAAACACGAAA
TTGTATTGCAATACTACATGAAGGACAACAAACTGAAATCCTTGAAAAGGG
TCCCACGATTGAACTTAAGGAATCAGAGGAATTCAAGTCACATCTATTAAA
ACTTTTCAAAGAAACTGATGTGGCTGTTATGTCTGGTAGTCTGCCCAAAGG
ACTTAATACTGATTATTATGCGGATATTGTGAGATTAGCAAAAGAACAAGG
AATTTTGACCATTTTAGATAGCTCTGGTCAATCACTTGAGGAAGTTCTTAT
TAGTAATGTGAAACCTACAGTAATTAAACCCAATATAGATGAATTATCACA
ACTTTTAAATTACAAAGTAACCAATGATATTAAAGAATTGAAAGCGGCAGT
AAGTCAGCCAATATTTAATGATATTGAATGGATTATTGTTTCATTGGGCAG
TGAAGGTGCTTTTGCAAAACATAATCAAAAATTTTATAAGGTGAATATTCC
CAACATTAAAGTAGTTAATCCTGTTGGGTCAGGAGATTCCACTGTAGCAGG
AATTGCTTCTGGACTCATTCATCAACAAACCGATGAAGAGTTATTAAAAAA
AGCAAATGCATTCGGAATGCTAAATGCAATGGAACAACAAACAGGTCATAT
TAATACAGATAAATTTGACGAAATATTCAAACAAATAGAAGTTATAGAGGT
GTAATTTATGACAAAATCACAACAAAAAGTGTCATCAATTGAGAAATTAAG
TAATCAAGAAGGTATTATTTCAGCTTTAGCATTTGATCAACGTGGTGCATT
AAAAAGAATGATGGCAGAACATCAATCTGAAACACCAACAGTTGAACAAAT
AGAACAATTAAAAGTACTTGTTTCTGAAGAATTAACTCAATATGCGTCTTC
AATTTTATTAGATCCAGAATATGGTTTACCAGCATCAGATGCTCGAAATAA
TGACTGCGGACTATTACTTGCATACGAAAAAACTGGATATGATGTGAATGC
GAAAGGTCGTTTGCCAGATTGCTTGGTAGAATGGTCTGCGAAACGTTTGAA

AGAGCAAGGGGCCAATGCAGTTAAATTTTTACTTTATTATGATGTAGATGA
CACAGAAGAAATTAACATACAAAAGAAAGCATATATTGAACGAATTGGTTC
AGAATGTGTTGCCGAAGATATTCCTTTCTTCTTGGAAGTTTTAACATATGA
CGACAATATTCCTGACAATAAAAGTGCAGAATTCGCTAAAGTTAAGCCACG

5 TAAAGTTAATGAAGCAATGAAGTTATTCTCTGAAGATCGTTTTAATGTGGA
TGTACTTAAAGTTGAAGTACCTGTGAATATGAATTTTGTGGAAGGATTTTC
AGAAGGAGAAGTTGTTTATACTAAAGAAGAAGCTGCACAACATTTCCGTGA
TCAAGATGCAGCTACTCACTTACCATATATTTATTTAAGTGCAGGTGTATC
AGCAGAATTGTTCCAAGATACATTAAAATTTGCGCATGATTCTGGTGCGCA

10 ATTCAATGGTGTTTTATGTGGACGTGCCACATGGTCAGGAGCAGTTAAGGT
ATACATTGAAGAAGGAGAGCAAGCTGCCAGAGAATGGTTGCGTACGGTAGG
ATTTAAGAATATTGATGATTTGAATACAGTATTGAAAACAACAGCTACATC
ATGGAAAAACAAATAATGTAAGGGAGGATATTCAAATGAATAGAGATGAGG
TACAATTACTCGGATTTGAAATTGTTGCCTATGCTGGGGATGCACGTTCAA

15 AATTATTAGAAGCTTTAAATGCTGCTAAAGATAGTGAATTTGATAAAGCAG
AACAACTTGTAGAGGAAGCGAATGAATGTATTGCTAATGCACATAAAGCAC
AAACCAATCTTCTAGCTCAAGAGGCTAAAGGCGAGGATATCGCATATAGTA
TCACTATGATTCATGGTCAAGACCATTTAATGACAACATTACTTTTAAAAG
ATTTAATGAAGCATTTAATTGAATTATACAAAAAAGGGAGCTGACCCCTAT

20 GAATAAATTAATAGCATGGATAGAAAAAGGAAAGCCATTCTTTGAAAAAAT
ATCACGAAATATTTATTTAAGAGCGATTCGTGATGGATTTATTGCTGCTAT
TCCAATTATCTTATTCTCAAGTATATTTATTTTAATTACCTATGTACCAAA
TGTGTTTGGTTTTACTTGGAGTAAAACTATGGAAGGTATATTGATGAAACC
CTATAACTATAC

25

Sequence 3415
step.1002g07.cons.ok
TTTCATCCTCAATATCCATACCAAGTAGCTCTTCTATTAAATCTTCGTGCG

30 ATACGATAGCATCCGTACCTCCAAATTCATCTAAAACAATCGCTAAATGTT
TTCGTGAAACGGTCATTTTACGCAATACCCATTCTGCCCTATTATGTTCAT
TTACAAATAAAGGGCTTGATGCATAATTAGTAATTGCGTCCTCTTTATTTT
TACTCCAAGCTAATAAATATTTTGAGTGGAATACGCCGATGATATCATCTA
TATTTTCATCATATACTGGATATCTTGTATATGGGTTATTCATCACTGTAT

35 CATAAGCTTCGTCATATGTTACTTCCTTTGAAAAAGCTACAACATTAATAC
GAGGCGTGGTATCAACATCCTTAACCTTCAATTGTTCAAAGTCCATAACGT
TTTGAAGTCGAGTATTTTCTATCTCATTAAATGCACCTTCTCTACCCGCAA
TATTTAATAATGTACGAATTTCTTCTTTAGAAAATCTTTTTTCAACAGGTT
GGCCTCGAGATAGTAAATGATTAATACCATCTGTCATCTTGTTTAATAAGA

40 TTGTAATGGGCTTGAGTACAATAACACATATATGAATGATAGGATACACAA
GTTTTGAAATTTTATCAGGAAATGTTGCAGCAATTGATTTAGGAATCACTT
CAGAAATAAGTATAATCACAATTGTTACAATTGCTGATGCTATTCCCACAT
TCACTCCAATATCAATCGCTAAGATTGTCACGAGTGTAGGTAAAATAATAT
TCGCTACATTGTTACCGATTAATATAGTAGTGATAAATTCACTTGGTTTAT

45 CTAATAACTTTGTTAAACCTTGCGCCTTACGATCTCCCTTTTTAGCTTCAG
TTTGAAACTTAGTGCGATTCGCGGCTGTTAATGCTGTCTCACTACCTGAAA
AGAAAAATGAAACAAATATTAATAGTATAATTGCAATAATCATGATGCATT
TCTCCTTAAGATTTAGTCCGTATTAATACATTATTTCCCGAGAAATGAATT
ATCAAACTTATCATTTATATTGAATTAACACAACACACGAATAAGAATGAA

50 CTAATGCAAAATATTTTTATAGTTTGCCTATAATTAATAAAGAAGATATCG
ATTTAATTATTAATAGAAGTTTAGAACAAAATGTATTTCTACTAACTGAAT
ACATTCAAAAAGGTAATAAGAATAAAGCTATACAATTAGTCAATGATTTAA
TCATTATGAAAGAAGAACCTATAAAATTACTGGCATTAATTACTAGTAATT
ATAGATTATATTATCAATGTAAAATTCTTAGTCAAAAAGGCTATAGTGGGC

55 AACAAATTGCAAAGACTGTAAATGCACACCCTTATAGAGTAAAACTAGCAC
TCAATCAATCCCGACATTATAAACTAGAAAGTTTGTTTAACATCATAAACG
CTTGTGCAGAGACTGACTACAAACTAAAATCATCTTATATGGATAAACAAC
TCATTTTAGAATTATTCATACTATCTCTATAAAAAATTTGAGGTAGGAGAT
ATAAAATAAATCCTATCGACAATAGAAATATTGGTTGTAAAGTAATAAAAA

AGAGCCTAGGACAATTCGTTTTGTCCTAAGCTCTGTTTTAATAATCATTAT
TTATTAGCTGACATAAGTTTAGATTTAATACGGTCAGCTTTATTAGAATGG
ATTAAATTACTTTGAGATGCTTTGTCAACTTGTTTGATAGCAAATCTTAAT
AATTCATCTTTGTTTTCAGCATCAGTAGAGATAGCTGTTTTAGCTCGTTTC
5 ACAGCTGTACGCATAGCGTTTTTCTTAGAAATATTTCTTTCTTCAGCTGTT
TCAGTTGTTCTTACACGTTTGATTGCAGATTTAATATTTGGCATTACTGTC
ACCTCCTAAAAGTGATCATAACTTATCAAAATTTATTTGATTACAACAAGA
AATATTTTATCAAATGTCTTATATTTGTGCAATCATTTATTTAAAAGTAAC
TTCATTTGAATTTATTACTTTTTAATTGTTGGAATGAACTACTGAACTAGA
10 TATAATAGTTATGTTAGAATTTACATTTAATTAATTATGTCAAAGATATAT
TGCATTAAAGTAATAAAAACGTTACTATATCAAAGATACATAGAAGTGACA
GGTTATAAAGATGAAAGCGAGAAGGATAAAATGGATAAGCAAGAACGATAC
AATAGAAGAGAAAATATTAGAAATTTCTCCATTATTGCTCATATAGACCAT
GGTAAATCGACATTAGCTGATCGAATTTTAGAGAATACAAAATCAGTTGAA
15 ACTCGAGAAATGCAAGATCAATTACTTGACTCTATGGATTTGGAAAGAGAA
CGAGGCATCACTATTAAACTAAATGCTGTTCGATTAAAATACGAAGCTAAA
GATGGAGAAACTTACACATTTCATTTGATAGATACACCAGGACATGTCGAC
TTTACATATGAGGTTTCTCGCTCATTAGCTGCATGTGAAGGTGCAATTCTT
GTAGTTGATGCTGCCCAAGGTATAGAAGCACAAACCTTAGCAAACGTTTAT
20 TTAGCATTAGATAACGATTTGGAACTTTTGCCAGTTGTTAATAAAATAGAC
TTGCCTGCAGCTGAGCCCGATAGAGTTAAGCAAGAATTAGAAGATGTTATA
GGTATAGATCAAGAAGATGTAGTACTTGCAAGTGCTAAGTCAAATATAGGT
ATTGAAGAAATTTTTAGAGAAAATAGTTGATGTTGTACCAGCACCGGACGGT
GATCCAGAAGCCCCACTTAAAGCACTTATCTTTGATTCAGAATATGATCCA
25 TACAGAGGAGTAATATCTTCAATTCGAATTATTGATGGTGTTGTTAAAGCT
GGAGATAGGATTAAAATGATGGCTACCGGTAAAGAATTTGAAGTTACAGAA
GTCGGAATCAATACCCCTAAGCAACTACCGGTAGAAGAATTAACAGTTGGT
GATGTGGGTTATATTATCGCAAGTATCAAAAATGTTGATGATTCTAGAGTA
GGTGACACAATTACTTTAGCTGAAAGACCTGCTGACAAACCGTTACAAGGA
30 TATAAAAAGATGAATCCAATGGTATTTTGTGGTCTATTCCCTATTGACAAT
AAAGACTATAATGACCTAAGAGAAGCTTTAGAAAAATTACAA

Sequence 3416
35 step.1002g10.cons.ok
TGTAATACCAATAAGTTCAGTTCCTTGGAATGAGAAGCCTGCGACTAGAAA
TACACCAAGAATTGATAGCAAGCTACCTCCTAAATTCCCACCTAGTATTGG
CCCATCTCCTTTTGTAAACGTATCAAATCCTACAAATTCTCCACCCATAAT
CCCTAAAATAGTTAAGATACCAATTCCTATAAATATGATGACTGTAACTAC
40 TTTGATAAGTGCGAACCAATATTCACTCTCTCCATATACTCTAACGGATAA
AGAATTAAGCGCGAAAATAATAATAAGGAAAATACAACTCCAGACCCAAGC
TGGTATACCTTGCATAGGGGACCAATATTGTATAACTTGCGCTGCAATAGT
AACATCAGCTGCTACTGTTATCACCCAGTTGAACCAATAATTCCATCCTAA
AGCAAAACCTAGTGAAGGATCGACAAAGCGTGTAGCATAAGTACTAAATGA
45 ACCTGACACAGGCAAATATGTCGCCATCTCCCCTAACGACGTCATTAGAAA
GAACACCATCGCTCCAATAACTGCATAGGCAATCAAAGCACCCAATGCACC
TGCGTCATGTATAGCTCCACCAGAAGTCATGAATAGACCAGTTCCTATACA
TCCCCCAATGGCAATCATTGAAATATGTCTATCTTTTAATCCTCTTTGAAC
ACCGTCATTTGTTTATTATTCTTTTTCATAGAAGCTCCCTTTCCTTCATA
50 AGTAGAGGCTAACTTGTTACCACATCAATGCATATCATTGTTGTAGATAAA
TTTGGTATAGATGATATAAAAGTAAAATAATATTAATATCATTTTACTTTT
ATATCACTTGTCTTTATATCTGTATCAATAATGTAAAAAGTGGTTACATAC
CTTTAAGATATGCAACCACTAATGTTTATCTATATAAGGTAGCACATCACA
TAAAGTGACAGTACAACTTCTATTCGAAAGTTGTCCCAACAATGCTAATTT
55 ATGGAAATAACATCATTTCGGCATTATTACCTTTCACTTATTCAACACATG
TCCCATACGACTTCTGATAAGTTAGTCATTAATAATGCAACCTCTAACTCT
ATTTATTTATATTAATTTGTAATATACATGATATACTTTATATTTTCAAGTT
TATCTATGCTTTTTCAAACGGATTTCATCAATAACATTCCAAATAAATTCA
TCTTTTTCAACTGATTCTTGTTTTTTCAGAACCATATTTTCTTACATTTACT

TCTTGATTCTCTACTTCTTGGTCACCTACAACAATCTGATAAGGTATTTTT
TTCATTTGAGCTTCACGAATTTTATATCCCATTTTTTCATTACGGTCATCA
ATTTCAACGCGGACACCTTGCGATTTTAGTTCATCTTGTAAAAGTCTTGCA
TAATCATAATGTAAATCTATATTTACAGGAATAATTTCAACTTGCATAGGC
5 GCCAACCAAGTTGGAAATGCACCTTTTGTTTCTTCTGTTAAAAATGCAACA
AAACGTTCCATAGTAGAAACTACACCACGGTGTATAACTACAGGACGATGT
TGTTCTCCATCTTGACCAATGTACGTTAAGTCAAAACGTTCTGGTAAAAGA
AAATCAAGTTGTGCTGTTGATAGAGTTTCTTCTTTTCCCATAGCTGTTTTT
ACTTGAACATCTAACTTAGGTCCATAGAATGCTGCCTCACCAATAGCTTCT
10 TCATAAGTTAAACCTAATTCATCTGATGCTTCTTTAAGCATGGATTCAGCT
TTTTCCCACATTTCATCATCATCAAAGTACTTATGCTTATCTTCAGGATCT
CTATAACTCAATCTGAAGCGATAATCTTCAAAACCAAAATCTTTGTACACA
TCTTGAATCATATTAACTACACGTTTAAATTCTTCTTTAATTTGATCAGGT
CTAACGAAAATATGGGAATCATTCAATGTCATTCCTCGAACACGTTGTAAA
15 CCTGATACTGCACCACTTGCTTCGTAACGATGCATAGTACCCAATTCAGCA
ATACGTATAGGTAATTCGCGATAAGAATGAGGTTTGTTTTTATAAATCATC
ATATGATGTGGACAGTTCATTGGTCTTAAGACCATTGCTTCGTCTTCATCT
AACTTCATTGCTGGGAACATATCTTCTTGATAATGATCCCAGTGACCAGAT
GTTTTATATAAATCAACATTGGCTAATACTGGTGTGTAAACATGATCGTAT
20 CCCATACTTACTTCTTTATCGACAATATAACGTTCTATTTCCCTACGTATT
GTAGCACCATTTGGTAACCATAATGGTAAACCAGCACCAACGAGTTGATTG
TTTGTAAACAATTCTAAATCTTTACCAATTTTACGATGATCACGCTCACGA
CGTTCTTCCAACATTTTTAGATGTGCTTTCAAATCTTTTTTTGTCAAAGAAT
GCTGTACCATAAATTCGTTGTAACATTTTATTATCACTATTTCCACGCCAA
25 TAAGCACCAGCTGTAGATAATAGTTTGAACTCTTTAATTTTAGAAGTAGAA
GGTACGTGTACACCTCGACATAAATCAGTAAATTCACCTTGAGTATAAAGT
GTTACACTCTCATCTTCAGGAATTGCATCAATAAGTTCTAATTTATAAGGG
TCATCCTTGAAGAAGTCTTTTGCTTTTTCTTTACTAACTACTTCTCTTACA
ATTTTATGATTTTCGTTCACAATTTGTTTCATTGTTTTCTCAATTTTATCA
30 AAATCATCCGATGAAACCTTATCATCCATATCAAAATCATAATAGAATCCG
CCTTCTATTACAGGTCCAACTCCAAATTTAACGTCTCCGTATAAACGTTTT
AATGCTTGTGCCATTAAATGAGCTGTTGAATGACGAAGTACTTCTAACGCT
TCTTCACTCCCAGGAGTAATAATTTCAATAGCTCCATCTTGTTCTAAAGGG
CGTGTTAAATCTACAAGTTGACCATTGAATTTTCCCGCAACTGCTTTTTTT
35 CTTAATCCTGGACTAATTGATTGAGCGATGTCTTCTGTAGTAGT

Sequence 3417
step.1002g12.cons.ok
40 ATACTGCGATATTAACCGATGAAGAAATACGTGAAAAAACAAAATCTTTTC
AAAAAGAGTTAGCAGAAATTGAAGATGTAAAAAAACAAAATGATTATTTAG
ATAAAATTTTACCTGAAGCTTATGCACTTGTACGTGAGGGGTCAAAGAGAG
TATTTAATATGATTCCTTATAAAGTACAAGTAATGGGTGGTATTGCTATAC
ATAAAGGTGATATTGCAGAAATGAGAACAGGTGAAGGGAAAACATTGACTG
45 CAACCATGCCGACGTATTTGAATGCTTTAGCTGGTAGAGGTGTACATGTTA
TTACAGTCAATGAATATCTATCAAGTTCACAAAGTGAAGAAATGGCTGAAC
TATATAACTATCTTGGCTTAACTGTAGGTTTGAACTTAAATAGTAAGTCAA
CTGAAGAAAAACGTGAGGCTTACGCACAAGATATCACTTATAGTACGAATA
ATGAACTTGGGTTTGATTATCTTAGAGATAATATGGTGAACTATGCTGAAG
50 AGAGAGTAATGCGTCCTCTACATTTTGCAATTATTGATGAGGTCGATTCCA
TATTGATCGACGAAGCAAGAACACCTTTAATTATTTCTGGTGAAGCGGAAA
AATCTACTTCTTTATATACTCAAGCAAATGTTTTTGCAAAAATGCTTAAAG
CGGAAGATGATTATAATTATGATGAAAAAACCAAAGCTGTACATCTTACAG
AACAAGGTGCAGATAAAGCTGAACGTATGTTCAAAGTAGATAATCTTTATG
55 ATGTTCAAAATGTGGAAGTGATTAGTCATATTAATACAGCTTTAAGAGCTC
ATGTTACTTTGCAACGCGATGTTGATTACATGGTCGTTGACGGTGAAGTAT
TAATTGTTGACCAATTTACTGGACGTACAATGCCTGGACGTCGTTTTTCTG
AAGGTTTACACCAAGCAATTGAGGCTAAAGAAGGTGTAGCAATTCAAAATG
AGTCTAAAACGATGGCATCCATTACTTTCCAAAACTATTTCAGAATGTATA

ATAAGTTAGCGGGGATGACTGGTACAGCGAAAACCGAAGAGGAAGAATTTC
GTAATATCTATAATATGACAGTTACCCAAATTCCAACAAACAAACCTGTTC
AACGTAAAGATAATTCAGACTTAATTTATATTAGTCAAAAAGGAAAGTTTG
ATGCGGTAGTTGAAGATGTTGTAGAAAAACATAAAAAAGGACAACCCGTCT
5   TACTAGGTACTGTTGCTGTTGAGACTTCTGAATATATTTCAAATTTACTAA
AAAAACGTGGTGTCAGACATGACGTATTAAACGCTAAAAATCATGAACGCG
AAGCTGAAATCGTTTCAAACGCGGGGCAAAAAGGTGCAGTTACAATTGCCA
CAAATATGGCTGGACGTGGAACAGATATTAAACTTGGTGATGGTGTTGAAG
AGTTAGGTGGACTTGCTGTTATTGGTACTGAGCGTCATGAATCAAGACGTA
10   TTGATGATCAATTACGTGGACGTTCAGGACGCCAAGGTGATAGAGGAGATA
GTCGTTTTTACCTATCTTTACAAGATGAATTAATGGTACGTTTTGGTTCAG
AACGCTTACAGAAAATGATGAACCGTTTAGGAATGGATGATTCAACGCCAA
TCGAGTCGAAAATGGTATCTCGAGCTGTAGAATCAGCTCAAAAACGAGTAG
AAGGTAATAACTTTGACGCGCGTAAACGTATTCTAGAATACGATGAAGTTT
15   TACGTAAGCAACGTGAAATTATTTATAATGAGCGTAATGAAATCATTGATA
GTGAAGAAAGTTCTCAAGTCGTTAACGCGATGTTACGTTCTACATTGCAAC
GTGCGATTAATCATTTTATTAATGAAGAAGACGATAATCCTGACTACACGC
CATTTATCAATTACGTTAATGATGTGTTCTTGCTGAATTATTAATGTTAAA
TGCTATCCATGTGCCATTACACACCGGTAAAATTGAAGAAATGACACGTGT
20   ACTTCGTGAAAAATAAGCATAATATATTTAAGCGTCTGGATTATCATCCCC
AGACGCTTTTAAAATGTTTAGAGAATTATCAAATTTTATTTGAATGCTTTC
GCTAAAACAAAGATAAACATATCTTTCGAGGAGTTTTATCTAGTATACGTA
CATTCTTCCCTCAACAAGCTAAAAACATAACTATCTGAGAATTCCCCTTGT
AATCGTTCATTACTTCTCAATATACCTTTGCGTGAACCCAAGTCTAATTGG
25   TATAGCTCTGCTTTTCTTATTTTTAGTAGACATTCGTATTTCTATACGATT
AATATCATACACTTCATATACATAGCGAATTAACGCTTTAGTACATTTAGT
CATAATACCTTTCTTTTGAAAGTCTTCAGCTAAATAATAACCAATTGAAGT
TGTTTTATTAACTAAATCTAAGTAATGCAATCCTATGACTCCAATCAATTC
TTTATTACTCCATATTCCACAATGAAATCCATTACCATCGATAAATTGTTG
30   CAACGCCGAATGGATAAAGTGTTTACTATCTTCAACTTTCTTCGTATGTTC
AACAAAAGGCAGAAATTCAGCTAAATAGTCACGATTGCTATCTACTAATTT
AAATAACTGTTCGGCTTCTCGTTCTTCTAATATTTTTAATTTAATATTTTT
ATTTATTTGATAACTAAACATACAAAATCATCCTTTCTTAGAATTTTCTGT
CATTTTAGTATACGCGAGTATAGGATTATGTCATAAAACTTGATTACTAGA
35   ATTCGATATTCTTTAAGTTACAAGTGTGTATGTAGTATTTCTCTTGCTTCT
ATTGTCTTATACTTACAACCAATATTAGAGTGAAAAAGCTTAACTTAAGAT
AGATAATCATCAACAACACAAACGTTAAAAACACTGAGAAAACTTTTTTAA
AATATACTTCTCAATATTGATTAGTAAAATCCAAAATATGAATAATAAGTA
GACATTAAACAATCTTTATTTTATTCTATAAATAATATAAAATAAATTATT
40   AGATTAATACATTATTTTTTTATTTCTAATTTAAGTTGAAATCGTATGTAA
TCGGAGGAAACCATAATGATATATTCTATCGGTAAGAATTTAGGTAATAAA
TTAACAGGTATAGAACAAGCTATGATCAATAGATTAAAGCTATTTAAAGAT
AATTTAGTCCCAAATAAACTCATATTCACATCTTGGTCACCACGTTATAT
ATGCATGCACATTCGTTAAACATCGATTCAAAAGATATTTTCAGTCTTTAC
45   GATTTTCTACAAGATAGTATTAACTTTGAGAAAAAACATATTGATTGGATA
AATTATTGGCAAAATATATGTAATTATACCTTAAAATTCGTTGAAAATACG
AATGATATTAAAAATATACGATAACGACACATATAAAATGTATGTGCATTTT
GTTGATTCAAATTATCAAACTTTAGACTATATTAACCATTTTGATATACAA
CAACGTAAAATTCGAAGAGATTTTTACGATACAAGAGGCTTTTTAAGTTGT
50   AGTAGAATTTTAACCTCTCAACAAAAGTCGTGATGGAACAATTTTTTACA
CCTACACAAAAAGTTAAATTTCAAAAATATTACAACCCTGAGCACGAACAT
CCTACGGTACAATCTATCATTT

55   Sequence 3418
        step.1002h01.cons.ok
        CTTGAAATATCATTTGAATTTTCTTATTAAACTTCAATCGCTCTTTACGTT
        TTCTAATTTGATGAATATCTTTACCATCATACTCGATTGCGCCATCAGTAA
        TATCATTTAATTTGATAATGGCTTTTCCAGTCGTTGATTTACCACATCCAG

```
       ACTCTCCTACTAGGCCTAAAGTTTCCCCTTTATAAACATCAAAACTTATAT
       TATCAATCGCTTTAACTTCATTATGTTTACCTTGATTAAAATACTGTTTTA
       AATTTTTAACTTGTAATAATACATTTTCATTACTCATTAAAAGCCACCCTC
       TCTACTCTATGTGGTTGTTCATAATTACTTGGCATTGTTCTTAATCGTTTT
  5    TGAACCATTGAAGGTGGCGTAACTTTTGGAGCTCTTGCATCTAATAACCAA
       GATTTAACAAAATGTGTGGGTGAAATTTTGAACCAAGGTGGTTCTTCTTTA
       AAATCAATATCTAAAGCATATCGACTTCTACGTGCGAAAGCATCACCAATT
       GGTGGATGAAGTAAATCTGGTGGTGTACCTGGAATTGCAATTAAGTCCGTG
       TCATTACTGGTTGTTAAATCAGGCATTGAAGAAAGCAATCCCCAGGTATAG
  10   GGATGTTTAGGATCATAAAATATTTCATTCACATCCCCTGTTTCAACCATC
       TGTCCCCCATACATTACGGCTACTTTGTCCGCAATATTTGCTACAACGCCT
       AAATCATGCGTAATAAAGATAATTGAAGTTTCAATCTTATTTTGTAGTTCT
       TTCATAAGATCTAAAATTTGAGCCTGCATTGTCACATCTAAAGCAGTTGTA
       GGCTCATCAGCAATTAATATTTTAGGCTCACATGCCAATGCTATTGCAATA
  15   ACTATTCTCTGACGTTGTCCTCCTGAAAATTGATGTGGATAAGCTTTAAAT
       CGTTTTTCAGCACGAGGTAACCCAACTAAATTCAAAATTTCCAATGCTCTT
       TGCTTGGCCTTTGCTTTACTTAATTTCTTATGTTTAATCAAAGGTTCCATG
       ACTTGCTTTCCGATTTGCATTGTTGGATTTAAAGAAGTCATAGGATCCTGA
       AATATCATTGAAATATCTCGACCTCTTAGCTGTATCAGTTCTTTTTTCACTT
  20   TTCTGAGCTAAGTCCTCACCTAAAAATAAGATTTCTCCCTTTTTTATTCTT
       CCTGTATCCTTTTGAAAAAGTTTTGTAATTGCCTTAGTTGTTACAGATTTT
       CCAGATCCAGATTCTCCAACAATGGCTAACGTTTCCCCTTTATTTAAATGA
       AAATCCACGCCTCTGACAGCTTGCACTTCTCCTGCAGCAATATCAAAGGAA
       ACGTGCAAATCATTTACTTCTAATACCGTTTCTGTCATTACTTTTACCTCC
  25   TTTATTTTCGCATTTTTGGGACGTTTAAAACGTTTAGCTTTCTTTCTAATA
       CTTTCGTTTTTACTTTGTTTAGCAGCGTCAAAGATTAAGCTGTATAAACCA
       GATTCGTTGATGATTACCATGTTTCGTTTTTGACCTGATGCACTAACTTGG
       TGCGTCAGCTTATCTTCATCATCAACATGATTTCTAATTGCATTATCAGAT
       CTTGTGTAACCTAATATTTCTGCCACGTCTTTACCTAAAAAATATGGTTCT
  30   CCATCTACTTCTATTTTTCTTACTGGTAAATCTTCAAAATTAAATACTTGT
       AAATCTTGCATATTGTTTATGCTCCTTTCTGCTATACTCCTTATAAAAGGA
       GGTGAAAATATGAAAAGTTTTATTATTGCGTATGATTTAAATAACCAAAAG
       GATTATCCAAAATTAATAGAGCGTATTGAGGATTATCCTAATGTTGCTAAA
       ATCAATAAATCAGTTTGGTTTATTAATTCAACTAATGATGCTAAAACTATT
  35   AGAAACGAATTAAAAATGTTTATTGATAGCGATGATAGTTTGTTCGTTGGT
       AAGCTGACTGGTGAAGCCGCATGGTCTAATGTAATTTGCAGTTCACAACAT
       TTAAAAGATTATCTTTAGTATCGTTAATCTTGAGCCTCTCTTTTTGAGGGG
       CTTATTATAATTTCGTCAAATTCTTCTGCATTTTTTATATTAGCTGCGACT
       TCTTCTAAACTTTCAAAAACTAATGTATATCCACTATTCTTTGTTTCTTTT
  40   ACAAAAACCCTGTATAATACCTTGTTTTCAGTCATTTGAATTCCTCCTTTA
       AGTTAAAACTTTCTTTTTGCGTAAGTCTTTGTTAAAAAAAAATATCTCTTCC
       TTCTTGAGGTGTTAAATCTAACGCGAAATAAATTCCATTTATTACTGGATA
       TGAAGGTTTAGTTCTTCCGTGTATCATATTGGACAATGTATCTCTATTAAC
       ACCTATTTCTTCAGAAAGAGTTTTGATGTTGTGTCCTTTTAAAGCCATTTT
  45   TGATTTTAAAAGTTTAGTGTCGATAGGCATTTCATTTCACCACCTTTCGTA
       TTACGTAAGTAATCTTATCATGGCTGTACAAAATAGGTCAAGCATTTTACG
       AAAGTTTTTAAGAAAAAATATTGCAAATCACGAAAGTTTTCCTTATAATAT
       AGTTATAAAGTAAAAGGAGCTGTATTACGATGTGCTTTTCAAAAAGAATGA
       AACAATCAAGAGAAAAACAAGGTATGACTTTAGCTGAACTAGGAAGAAAAA
  50   TCGGTAAAACTGAAGCTACTGTACAACGTTATGAAAGCGGGAATATTAAAA
       ATCTTAAAAATGATACTATTGAAAGTATAGCTACTGCATTAAATGTTAACC
       CTGCTTTCTTGATGGGTTGGATAGAAGAAGTTGAGGAACAACCACAACATC
       GTGCAGCGCATCTTGATGGTGATTTAACTGACGAAGAATGGCAAGAAATTC
       TTGATTACGCTGAATACATAAGAAGTAAAGAAAATAAAGGGTGTTTTATG
  55   TGGGGAAATATGAAGATATGTTAATTGAACATGACTATATTGAAGTCATTG
       AATGTGATAACTTACCTAAAAGGTTATCTGGTTTGTGGCTTGGAGATATGA
       TTTTAATTAATCGTAACTTGCCTATTACTTCCAAACTTGAAACACTTGCAG
       AGGAACTCGCTCATAACGAACTTACATATGGAAATATAGTTGATCAAAGTA
       GTTTTAATCATAGAAAATTTGAAGGTTATGCACGTAGGTTAGCCTATGAAA
```

AGTTAATCCCTCTTAAAGATATTGTAAAAGCATTTTTGCAAGGCATTCATG
ACTTGTATGAACTTGCTAATTTTTTTGA

5   Sequence 3419
       step.1002h02.cons.ok
       AAATAAACGAGAAAACCATAATCATGATAATCAACCATACAAATAGTACAA
       ATGTATTAAAGGTTGGATACTTATATCATATGTTTTAATGTTTGAAACTTT
       TATTATAAATAAATGAATTGAAAAATGCTTGTTGCTGACTGTTTCGAATCA
10   TAGTCAGTTGTAAGGGAGCGAGACGACGAATCTGAAAAGAATCGGTCTCGC
       TTTTTTAGTATAAAAATTTAATTATTATTTTAAAATGATAAAGTGATTTCT
       ATAGTTAAATTTAGTGCGAGGATACTTTAAAAAATTATATTGAAAAAATGT
       CGCTATTACAAAATGTTCTAAAGTGCACAATTCAATTTAAAAACAAAAAAG
       AAGTTTTAAAATTTATATTTCTGTCTATATATGAGTTGTGCACGAAATGAC
15   AAACGCAATTCACATATGAAATATTTATCACTTTACAATAATATGTTGTTA
       AGTTGATGTAGATAAATTCATTGTTTTACGTATAAGTCACTATCATTGTTT
       TTTAGGATTCATTTTAATTATATCCAATCGACTAAAGTGAAAAGGAGGAAC
       ATATTAAATGAATTGGATATCGATTATTTTATTTATTATGGTCGTATGTGG
       CATTTCTTTTTATGCCTATTTGCAATCAAGAAAAATTAAAACGAATAGTTC
20   AGATGGTTATTTTATGGGAGGAAATAGCCTTACTGGATTTACAGTTGCCTC
       TACAATTATCATGACCAATTTGTCGACAGAACAAATTGTTGGGCAAAATGG
       TCAAAGCTATGCACAAGGAATGGAAGTAATGGCTTGGGAGGTTACGGCTGC
       CGTAGCAGTTGTATTGTTGGCCTGGGTCTTTCTTCCTAAATATCTAAAATA
       CGGTGTTAATACAATTTCTGAGTTTTTGGAATTACGTTATGATACATTTAC
25   TAAACGTTTTGTCTCCATTTTATTTATCTTTACCTATGTCGTATCTTTCTT
       GCCTGTAGTATTGTATTCTGGTTCCCTAGTGTTCAATAAAATGTTTAAAGT
       TGATGAATACTTAGGTGTAAGTAGTTCAACTGCTGTCATTATTATTTCATC
       TATTATTGGTATAATTGGCATTATTTACTTATTTATAGGTGGTTTATCGTT
       AAGTGCTTTTAGTGATTCGATTTATGGCATGGCTTTAATTATAGCTGGACT
30   TGCGATTACAATATTAGGTCTAGGTCAATTAGGAGATGGCAACTTCCTACA
       TGGTTTCGACAAAATCGTGCAAGACACGCCTGAGAAATTGAATGGTTTTGG
       TAAGGTGGACTCGGATGTTGTACCTTGGCCAACCCTATTCTTCGGTATGTT
       CTTTAACAATTTATTCTTCTGGTGCGCAAACCAGATGATAGTTCAAAAAGC
       ACTCGCAGCTAAAAATTTAAAAGAATCTCAAAAAGGTGCAATATATTTAAG
35   CTTATTTAAAGTGTTCGGACCTTTATTTACAGTCTTACCAGGCGTAGTAGC
       ATTTAACTATTTTAATGGTAGTCTTGATAAATCAGATAACGCTTACCCTGC
       ACTTGTAACTTCAGTATTACCAGAATGGGCATTCGGCTTATTTGGTGCGGT
       TATTTTTGGTGCAATATTGAGCTCATTTGTTGGCTCATTGAATAGTACAAC
       TACACTATTAACACTCGATTTCTATAAACCTATTTTTGGAAAAAATAAATC
40   AGATAAACATATTGCTCGAGTAGGCCATATTGCTACTGTAGTCATTGGAGT
       TATTGTTGTAGCACTTGCACCAGTCATCTCATTATTCCCTAGTGGTCTTTA
       TGCAGTAGTTCAACAGTTTAATGGTGTGTATAGTATGCCAGTGCTAGCTTT
       AATTTTAATCGCTTTCTTTTCTAAACGCACATCTAAATTGGGCGCTAAAGT
       GACACTCTTCACACATATAATTTTATACGCTATAATCAGCTTTGTATTTAC
45   AGAAATTAATTATCTATACACATTTAGTGTATTATTCTTTGTAGATTTAAT
       TATTATTTTGATTTTTAACAAAGTTAAACTATCTAGTGAGTTTGATTTAAG
       CACGCACCAACCGAAAGTAGATATGACGCCATGGAAATATCGTTACGTTGC
       AGGTATTATTGTTCTTGCTTTAGTAGTAGTAAGTTATATTATCTTCTCACC
       ACTCGTGTTGGCAAAATAAATGTTACAATTAAAACTTAATGAGTAAGAAAA
50   TAAAAGTGTACGTTAAGGACGTTTATCAGCGTGATCACGTGCACTTTAACT
       ATGAATAAAGTTGATTTAAAAATTAGAAAGGAGGGACAACGACGATGCAAA
       CAGTCGGAATTATACCTTCGCCAGGTATAGCACATCAACATGCAAAAAAAA
       TAATTCCAAATGTTAAACAGTTATTGTCAAAGCGTACTAAACATAGTCAAT
       GGAATTTCGACATCAAAGTCGATCTCATGATAGGATCTGCAGAGGATGTAC
55   ATGAAAGTGTTGAAAAAGCAGCACAAATTAAAGAGGAACATCAGTGGGATT
       ACGTTGTTTGTCTGACAGATTTGCCTAGTATTTCAGATAATAAAGTGGTTG
       TCAGCGACTTTAATAGTGACAAACATGTTGCAATGCTATCATTACCGTCAC
       TAGGTTTTATTGATTTGAAGCGCAAGCTAGTTAAAACGATGACTTCATTGA
       TTGAACAATTATATTATAATCAACCGAAAGACAAAAATGCGCCACATCCTT

934

```
TTGTACGCGTGAAGGCTGTAGAACCTGACGAAGACGCCACATCAAAACAAC
GATATATTAATATTTTATTTATCATAAGTTGGATTCAGTTAATTGGTGGAC
TGACACGAGCAAATCAGCCTTGGAAAAACATCTTTAATTTTAAGAAAATCA
TTTCAGTTGCCTTTGCAACAGGAACTTATGTCTCAATATTTTCAATGCCAT
GGGAATTAAGCGTGATTTATTCACCGCTTCGACTTATCATATTGATGGTGA
TTGCTATACTTGGGATGGCTGGATGGCTATTCTATGCGCATCAATTGATTG
AAAAGAAAACTGCTAAATCTCAGCGTGTATATCGATATATTTATAATTCAA
CCACACTTGTTACACTAAGTTTGATTACACTCATAAATTATGTCATTTTAT
ATTTATTGTTAATCATCAGTATTACACTCTTTGTCCCTGTGGAATTATTTA
ATAGTTGGACGAGTGCCCAATCACAATTTACGTTCTCAAATTATATGAGAT
TGATTTGGTTTGTATCATCATTAGGACTTTTAGCTGGAGCTATGGGATCAA
CTGTTGAAAATGAAGAGAAAATACGTCGTATTACTTATTCTTATAGACAAT
ATCATCGTTATAAAGAAGCTGGCAAGAACAAAAAGAACAAGAAACTTCTCG
TGATGTATCACAACAAAATGTCGAACAACAAACTTCAAGTAAAGATGAAAA
TAATGAACAATATGAAGGTAAAAAACAAGGACATAGAGAGGAGGATGACGC
ATGACAAATCAAAAAACTGTGGGTCTAGTCGTCGCTCCAGGTGTTACTGAA
CGCCTTGCAGAAAATCTCATACAAGAAATGCCTAAAATGTTATCTACGCAT
TATGATCATCAGCAAGAATGGATTTTTGATTTA
```

```
Sequence 3420
step.1002h05.cons.ok
CCAACGTGAGATAATCAGCACCTTTTTCATTAATAATAGCTGCGGCCGATT
CTAATATTAATTGTCTTTTGGAAGTTCTAGGCATTTCGTCACCTTCTTTAT
CGAATTCTAAATAATATATAATTGTAATTATTTAACACTTTTTCAACATCC
ATACTTATCGAGTGTATTTATTTTATAAACAACTTTTCTTAAATTCAACTT
TATTAACATGAATTGTATTTATTAATAACGATTTACTTATTACTATATGTC
ACAAACATATAAAAAGTAAACTTATCATATAATATATGTTCCTCTGTCAAC
GTCGGTTAATTATGAAAAAATATTAAGTAGCTCCATTCCGTTGCTTAATAC
TTTTTTATTCACTTATAAAAATCCAAGAAAAGAGAGAGCGATTTCGCCCTC
TCTTTTTTGTGTAGTAGTACTATTAAATTTTATCTCACTAACAATTAATAT
ACTTATTAATAAAATGATGTATTTAAAAATTATAATAATCTTGAAATATAA
TTATTAAGAAATTCGCAGTTATATCATTTATTAACTCAGCCTTATTGTACG
TGGGACGATTTTTTAGCAGTCATAAAGTATACAATAATACCAATAATGAAA
CCAATAAGTGCTGGAACAATCCATCCCATACCAATGTTAAAGAATGGGAGA
TAATTTTCACCAAATCCAATGATTGTTTGTGCAAATTTTGTATTAACAAAG
AACTCTGGACTTGCTTTCACACCATCTACAAATGCCGCCACCATTGTAAAT
AATGTTGTAAATCGATAGACAATTTTAGAATGTTTAAATAATGGACTAAGT
AATGCTAATAAAATTAAAGTAATCGCTAAAGGATAAATGAACATCAACACT
GGTGTTGAATACATAATAATTTTTGTTAAACCTACATTAGCAAATATACAA
GCTAATATACTCACCCCAGTAGCTAACCAAAGATAGTTAGATTTAGGGAAT
AACTCTGTAAATGTTTCCGAAAAGGCTGTGATCAATCCTATTGCTGTTTTC
AGACATGCCACAATGATGATTAGTGACAAAATAATAATTCCGTAATCCCCT
AAATAATGTTGTGCAATCTGAGCAAGCGCAATACCACCATTTTCACTTACT
TTAAAACGACCTAAACTCATCGTACCCATTAAAGCTAATAAAGTATAAATA
ACGCCCATAGCTATAATACTAATCGTACCTGATTTTAAAGTTTCTTTAGCG
ATTGTATTCGGATTAGTAATCCCCAACTTTTTAATTGTAGTAACAATGATA
ATACCAAATGCTAATGATGCCAAAGCGTCTAATGTATTATATCCATCGATA
AACCCTTTGAGTAACACGCTATTGCTATAATCAGCACTTACTGGCGCATGA
CTAATTCCACCCATAGGACGGATAAATGCAAGCACAACAACCATTCCAAGC
AAGATGAGAAAGACCGGATTTAAGAATTTTCCAATATAGTCTAATATTTTA
GAAGGTTTACGCGAAAATAACCACGCTACTCCGAAGAATAAAATACTAAAA
ATAGCAACAACGCTTGGGCCGTACCAGATGAAATAAATGGTGAAAATGCTA
TTTCAAATGACGTCGTCGCAAGTCTTGGCAACGCAAAAAACGGACCTATAA
CAAGATACAAGCCAATTGTGAACAAATAACCATATATTTTACTTATCCTTG
AGGAAATTTCAAAGACCCCGTTTGTTTTAGATACACCTATCGCTATAATTC
CTAAAAAAGGTAGTCCGATAGCCGTGATAAGAAATCCTAAATTGGCGGTCC
ATACATTTGCCCCCGCAGTTTGACCTAAGTGAATTGGAAAAATGAGATTTC
CAGCACCAAAAAAAAGACCAAACAGCATTGAGCCGATAAATAGATTCTCTT
```

```
TTAATGTTAATTTATTTTTCATCATTAATTATGTATCTCCTAATATGTATT
ATTAGCTATTATACATGATTTTAACCGTCTGCGTGGCACGAAGTTATATTT
ATATATGTGTGTATTATTACATCTTAATCTCATAGCTTTTCATTTTATAAC
ACGATTCAAACTATAAATAATGAAAAATTTTTAATATTTCTATACATTGTT
GTTTTTTATAAGAAATGTTAAGTTAATATTTATTTTTTATTAAAAATCCAT
TTATTTTAGCATATAATTTTTAATTAATAAAAGTTTTTGATAATATTAAGT
GTACAAGGATTTTTATTTTTTATTAAAAAGTGAAAATCATTACTAAAAATG
AGGAGGACTGTGAAATGACTGTATTGAGTGAGCAAGACAAAATAAGATTAT
TAGCTGATATCGTCAAGATTCAAACTGAGAATGATCATGAAATTAAAGTAT
GTGAGTACTTAAAAGACCTGTTAAGTCAATATGATATTGATTCTAAGATTG
TTAAAGTTAATGATTCACGCGCTAATTTAGTTGCTGAAATTGGTAGTGGCG
CTCCAGTCTTAGCAATCAGTGGACACATGGATGTTGTAGATGCTGGTGACC
ATGATGATTGGACTTTTCCTCCATTTGAACTCACAGATAAAGATGGCAAGT
TATTCGGCAGAGGTACCACTGACATGAAAGGTGGTCTTATGGCGATGGTCA
TCGCGATGATTGAATTAAAACAATCAAACGCATTAAAGCAAGGTACGATAC
GTTTATTAGCTACTACAGGAGAAGAGACGGAACAATATGGTGCACAATTAC
TAGCTGATGAAGGTTACTTAGATGATGTGAGTGGTTTAATTATTGGCGAAC
CAACTAGTAATATCGCTTACTATGCTCATAAAGGTTCTATGAGCTGTGTAG
TCACAGCGAAAGGTAAGGCTGCACATAGTTCCATGCCACACCTAGGTACAA
ATGCTGTTGATATTTTAGTTGATTTTGTAAATGAAATGAAACAAGAATATA
AAAATATTAAAGAACATGATAAAGTACACGAGTTAGACGCTGTTCCAATGA
TTGAGAAACATCTCCACAGAAAAATTGGTGAAGAAGAATCACATATCTACT
CTGGATTTGTAATGTTAAACTCTGTATTCAATGGTGGTAAACAAGTTAATT
CTGTTCCTCATAAAGCGACAGCTAAATATAATGTAAGAACTGTTCCAGAAT
ATGACAGTACTTTCGTGAAGGATTTATTTGAAAAAGTCATTCGTCATGTGG
GCGAAGATTATTTAACTGTAGATATACCTAGCAGTCACGATCCAGTGGCAA
GTGATCGTTGGAGATTTAATTAAACATTTTAAAAAATAGGAGTGCATTTAC
ATGTTTAAAAAGTTATTTGGAAAAGCTAAAGAAGTTGACAAAAACATTAAA
ATCTACGCACCTTTGACAGGTGAGTATGTCAAAATTGAAGATATTCCTGAT
CCAGTGTTTGCGCAAAAAATGATGGGAGAAGGCTTTGGTATTAATCCTACA
GAGGGTGAAGTTGTTTCACCGATTGAAGGCAAAGTCGATAATGTTTTTCCA
ACGAAACATGCCGTAGGGTTAAAAGCCGAAAATGGATTAGAGTTATTAGTT
CATATCGGATTAGATACCGTTCAATTGGATGGCAAAGGGTTTGAAGTACTT
GTTGAAAGTGGAGATGACATTAAAATAGGTGACCCACTTATACGTTTTGAC
TTAGAATATATTAATAATAATGCAAAATCTATTATTTCTCCTATTATTATT
ACAAACTCTGATCAAACTGAATCAATTCATATTGAAGACGTACAAGCAGTA
GTTAAAGGGGAAACACAAGTTATTGATGTGACAGTAAGCTAATGGTTAAGA
ATTATTCATTTTATCAATTTATTATGACCGTGCGCGGAAGAAAAGACGATA
AAGGGGTTTTTGCTGAGCAAATTTTTGAAGACCTTGCCTTTCCAAAACACG
AAGATGATTTTAATACATTATCTGAATATATTGAGACACATAGCGAATTCA
CCCTA
```

```
Sequence 3421
step.1002h06.cons.ok
ATACTTTCCTAAACTTCTAAATTGATTAGCATAAGCATCCAAATCGTCTTG
ATGTTTATTTAGTTCAATAATCTTGTCTGCAAATTGATTAATTTTAGGCAA
TGAATTATTAGCTTGATATACAGCACTCTTAATTTTATTGATTGTAGGTAC
ATTATCTTCAATTGATAGTCCGACTTTATTAGCTTCAGAAAGTAAAGCGGT
TGCTACGGTTTTGTTAAATTGTTTATTTGCTTTATCAATCACAAACGACGA
TCCCGTATCGGTTAACTTAGCTGCTACAGCATTAATCTTCTGATTTACTTT
AAAATCTATATCCGCCTTTTGAGGATGTTTTCTTAAAGTACCAGTGATTTC
ATGTGTGAATTTCTTCGGTATATAAATACCTGCATAATATTTTCCCATCTT
AATTTCATGGTCGGCTTTTTCTCTACTCACAAATTGCCAGTCAAAACTATC
ATTCTTTTTTAAAGTTTTGACCATTTTATTTCCTACATTAATATTCTTACC
ACGAACTTTTTCACCTTGGTCTTCATTCACTACGGCAACTTTGATATGTCC
TGTATGACCATACGGATCCCACATAGCCCAGAGGTTAAACCATGCATAGAA
TGAAGGAAGAATTGCTAAACCAGCTAAGATGACCCATACACCTGGTGTTTT
AGCAACTCTTTTTAAATCCGTGATAAAAAGTTTTAGTGCGTTTTTCATTTT
```

936

CACACTCCTAATTATCATTATTCTTTTTTATTGTAATACATTTCCGTACAG
ATGATTACTTAAATGTGCATATTTTTTTGCTTACGTAAAGATGATACACAT
ATTATATTAAAATTTGATTAAATGGATTTATTTTTCTAATGTTGATTAGTA
TAATATTTGTATTAAGGCGTAATTCATTTTTGTTGATGCTTGTCAACTTTA
5   TAAACCATAAGATAAATTATATCAAATTCTCATTTGTTATGACAATAACTA
AAAATAATTTTTAAGGGAGCACACTGTAAAAGCCTTTAGACCTTCCTTTCT
TAATGGCGTTTCAACATATCTAAATGATAATTCGGCTAATAATACAGTAAT
TACGATATCCATAATATAAACATAAGTTGGTAATTGACCATCAATAAAATA
ACTATGAATAAAGCTAATTACAGGAAAATGCCATAAGTATAAACTGTATGA
10   ACGCTTTCCAATGTAGACAAATAAAGGATTTCCTAATAACTTAGCTAAAAT
TGTCGTTGGATGAACAACACTTGCAATAATTAGCAAAGTCATTGTTGAAAT
AAGATAAAATCCACCGTTATAAATCCAATCACTTTCATCACTAACAGTAAA
GAATAATAGAATTAAAAATGTAAGTCCTATGATACCCGCACTATTAATCAC
AGTTTTTAATCCTTTAGGTGGATTGGGATTTAATTTAAAAGGTGGCCAGAT
15   AAATGCTAGAAGTACACCTAAAAGCAGTGTCTGCAATCTTGTATCAGTTCC
AAAATATACTCTAGAATGGTTCAAGTGAGGTTGAGAAATAACAACCATCAT
TAATAATGAAACCAGGGATATGATCCAAAACATCAGTATGACATTTTTCTT
TTTCTTAACTATTGCCATAAATAATAAGAGTACTGCTGGGAAAAAAAGGTA
AAACTGCTCTTCAATGGCTAGTGACCATAGGTGCTTTAAAGGCATAAAAGA
20   AAATTGCTCGAAATAATTGACATCTTTAGCAATATACCACCAATTAGATAC
GTAAAATATTGCTGCTATCATATCATGTTTAACTCTTACAATATGCTCGGG
GTGCAATAATAAAGTTGCAATTCCAACTACTACTATTAATGCAAATACCGC
TGGTAATAACCTTTTAATACGACGAATCCAAAAATTTTTAAGATTTATTGT
TCCAGTATCTTCATACTCTTTAAGTAATAAGCTCGTAATCAAATAACCTGA
25   AATAACAAAAAAGTATCTACGCCTAAAAAACCACCTGTTAACCATTGTTT
ATTCAAGTGATAAATAATAATACCAATGACTGCAATTGCTCGCAAACCATC
TAATCCAGGCATATATCGTATTTTTTTAGCAGAGCGATATTCATTATTTAA
TTTTTTTCTCATTGTTACATTCTTCCCTATTAATAATAAATTTCGTTAGAC
AAGTAACTTTTTTACTAATCAATTTCATAAGACGAACGCTATTGCTTCTAA
30   TCAACTTATTCAAACAAATTTTATCTCATTCAAATTAACAAGGTCATTCTC
TTATGTTATAGCTATCCCTATCATATAGCGCACGTTTCCATTGAGTTAATT
ATCATACTTCGTAAGATATACCTCATTAAAAGGTTATAAGTCTTAAGTATT
GTAATACATTTGTAATATTAATTCAATTTTATTATAGGCCTAATTTATAAA
AATATTAAATGAACATACGATAATTCTTATTGAATTTTCTTTAATAATTAA
35   AAAGGATAGCTACATCATTATTCCAAACAATGAGTAGCTATCCTTTATATT
ATCCGATGCCTAAAGCTATTTTGGCATATCGTGACATCATATCTTTATTCC
AAGGTGGACTCCATACAATCATCACTTCTGTATCAGAAATTTCAGGAATCT
CAGCCAAAACACTCTTAACTTGTTCAATAATTTGTGGTCCTAATGGACATC
CCATCGAAGTCAATGTCATTTCAACTGTACATAAACCTTCATCATCAACAT
40   CAACTTTATATACTAAACCTAAATTAACGATATCTATCCCTAACTCAGGAT
CTATTACCATTTCAAGAGCGCCTAAGATACTATCTTTTAGTGCTTCTTCCA
TCAGTATCACCTCTTTAAAAATTTTCTTTACACCAATATATCAAATATCCGA
CAAAACGCCAATAAAATGCTATGATACTGTTACATAATAATTAACAACACA
AGGAGTAATATATGCAACCTTATTTAATTTGTCTAGATCTAGATGGTACAT
45   TATTAAATGACAATAAAGAAATCTCACCTTACACTAAACAAGTATTAACCG
AATTACAACAATGTGGACACTACGTTATGATTGCTACTGGAAGACCCTATC
GCGCAAGCCAGATGTATTATCATGAACTAAATATGAGCACACCTGTTGTTA
ACTTTAATGGAGCATTTGTACATCATCCAAAAGCAAACGATTTTAAAGTGA
TACATGAAGTACTTGATGTGGAAATTTCTAAAAATATTATTACAGCACTTC
50   AACAATCTCATATTACAAATATCATTGCTGAAGTAAAAGACTATGTCTTTA
TAAATAGT

Sequence 3422
55   step.1002h07.cons.ok
ACATATCGCACGTGCAATAGCTACACGCTGTGCTTCTCCACCTGAAAGCAT
ATGTGGATAACAATACAAACAATCAGATTTTAACCCTACTTCACGTAGTAA
CGCTGTTACATACTCTTCCATATTCTCCTTTGATGTACATCGACACTGATT
CATTACTTCAAACAAAATATCTTTGACAGTATGAAAAGGATGCAAAGAAGA

AGTATAGTCTTGAAAAACCGCAGCAATTCGATGACGTCTCACTTTTTTTAA
ATGAACAGGCTGACCGTCCAACGTCACTTGTCCTTGATCTGGTTTTTCTAG
TCCTAAAATCATACGACTCAACGTAGACTTTCCACTTCCACTTTCTCCAAT
AATTGCTATTGATGCACCAGTGGGGCAATCAAATGACACGCCTTTAACTAC
AGGAATACGTTTATCACGAAATATTGATTTTTTTCCATAAAATTTAATAAC
ATTTTGAACCTTAAGCATTGAAATCACCCCTCAGCACTAACTGAAAATAAT
CATGAATTGTCTTTTTAGTAGATAATAAATAAGCTGTATAGTGATGACGAG
GTTGTTGTAATACCATATTTGTTCGACCTTCTTCTATCACGCGACCTTCTT
TCATAACTACAACACGATCTGCAATTTTATTAATCACTGTGAGGTCATGAG
ATATAAAAATCATGGCGCAATCAAAATGTTGTTTTATATCTTTAAAAGCAT
TGATGACTTCATATTGTGTAATGGTATCTAATGCTGTGGTAGGCTCATCAG
CAATAATCAATTTAGGTTTCAATGCCAATGCTAGTGCAATCATCATACGTT
GAAGCATCCCACCTGACAACATGTAGGGATATGCTTTTAATAACGAGCGGG
GATCTTTCAACTTCATATACGTCATATAATGAATGAGTTCTTGTTTTATTT
GTTGCTTTGATAATGTCGTATGTGCACGCAAAGTTTCAATCATTTGCTTCC
CCACTTTAGTTGAAGGATCAAATGCTCGTGTTCCTTGTTGCATAATCATAG
CAATATCTTTACCACGAATTTTGCGTAGTTGCTTTTCAGTTATTGTATTTA
AATTTTGATGTTCAAAAAAGATATCTCCGCTTACAGATAAACGTTGAGCAT
TTAAACCTATTAATGCTTTACAAGTAATCGATTTTCCGCTTCCACTTTCTC
CAATAATACCTAATGTTTCACCACATTGAACTGTGAAATTCACGTTATCAA
CGAGTGTCTCATCTGACCATGTGTCCTTAATTCTTAAATGCGCTACTTTAA
GCAATGTCATGTGCTTTCACACCTTTCTTCAATGCCAGTCGTTTCTCTTTA
GAGGACATACGAGGATCAATCGCCATTTGTAAAGCATCTGATAAAAAGTTA
AACGCCATTACAATTATGACGATAGCCACACCTGTTGTCATCATCATTCCA
GGATGTGTGAACATTACTTTTCGTGCTTCATTAAGCATCATCCCCCATTCG
GCTGTAGGTGCCTTAACACCTAATCCAAGGAACGAGAATCCTGACATTTGT
AATATCATTGAGCACATCGAACTACTAGCAATAATCGCTATGTCAGTAAAG
GTTAGTGGCAAAATATGTTTGCGAATGATTGTTAAATCATTCATACCAATT
ACTTTGGCAAATTTTACATGATCAGCTTCAATATATTGCATTACACTGGTT
CGAATCACGCGACAAAACCACGCCCATCGAGTCAATATAAACGCAATAATA
ATATTTTCTACACCCATGCCAAACAACGTAATCAATGCCAATGTCACCACA
TAGCTTGGAAAAGCTAACATCACATCGCATATACGCATAATTATTGCATCG
ATATAACCTGGGAAATAACCTGAAATAAACCCAAGTATCGCTCCTATCACA
ACGGAAATAATCAATGCGACAAATACATATAACAAACTAGGTCTTATGGCG
TATATTATCCGTGTTAATACGTCCCGGCCTAAATGGTCTGTCCCCAACCAG
TGAGACCAACTTATACCAGCAAATTTATTTGCTGTATCAATGTGATTCGCT
TCATAAAATGTAATTAAAGGAGCTAAAACTCCAAGCACTACGTAGATTGTA
ATAATAGCTATAGCAACTATGGCACCTCTATCTTGAAATAGACGACGTAAC
ACAATCATTTCATTGCCTCCCTTAAACGCGGATTTAATATTGCATTAATGA
TATCTGCAAGTGTATTAAAAAGAATAAACAACACCGCGACAATGAGCACGT
ATGCTTGAATTACTGGAAAATCATGTTCAGTAATCGCTTTAACACTTAATG
GTCCTAAACCTGGCCACGCAAATACATTTTCAATAACGACTAAACCTCCCA
TGATCATTGGTATCGACATACAAAAAATAGATACGGCTACCTGTAATGCAT
TACGTAATACA

Sequence 3423
step.1002h08.cons.ok
CTGGATGAACTTCAGCCATAACTCCATCAGCACCAACGGCTAAAGCCGCTT
TCGCTGTCGGTAACATAATATCTTTTCTACCTGTACTATGCGTTACATCAA
CCATAACAGGTAGATGTGTTCCTTGCTTTAAAATAGGTACAGCAGAGATAT
CTAATGTATTTCTAGTTGCTTTTTCGTATGTACGAATACCTCTCTCACATA
AAATAATATTTCGGTTACCTTGTGAAGCAATATATTCTGCTGCATAGATAA
ATTCTTCTATCGTTGCTGATAAACCACGTTTCAATAAAATAGGTTTATCTG
TACGACCAGCTTCTTTGAGTAACTCAAAATTTTGCATATTTCTAGCACCAA
TTTGGAAAACATCCAAATAATCTGAAGCGATTTCGAAATCATTTGGGTTAA
CTATTTCACTTACTACATTTAAGTTATATTTATCTTTAACATTTTTCAAAA
TTTTGAGACCTTCTACTCCAAGACCTTGGAAATCATAAGGAGAAGTTCTCG
GCTTAAACGCTCCGCCACGAATAAATTTCTCGCCTTTAGCCTGCAAATCTT

```
      GTGCTACAGCATCAACTTGTTCTTGTGATTCTACAGAACAAGGACCAAATA
      CAAATGATTTATTACCATCACCAATGATTCCACCATTATCAAATTTAACGA
      TCGTATCTTCAGGTTTCAATTTACGTGATACATAAAGATGTTTTTCATTTT
      CTGATTTTTGTAAATCTGTAGAGGCTTTAAAAATTTCTTTAAATAATTGCT ·
   5  TAATCACATTGTCGTTAAATGGACCTTTATTTTTATCTAATAGCTGATTAA
      TCATTTCCTTTTCACGCTGAGGATCATATACTCGTGTACCTTGCTTAATTT
      TTTCTTCACCAATTTTCTGTGCTAATTCTCCTCGCTTCGATAGTAATTCTA
      ATATTTGATCATTTAAGGAAACAATTTCTTCTCTATACTGTTCTAATTTAT
      TAGTCATTTCTACTCACCTCTAAGATATAGTTACGCATAAAGTATGTTACT
  10  TTAAAGCGATAAAATTAAAATTATTCAAATCTAATTTTAACAATTTCACAT
      TTATAGTTCAAGCGTCAATTTACAGAAAATACTAAATCTTCTGTAAATTTA
      TGTAAGTAATAATTTGTTAAATAATGATATTACAATATTATTTAAATGTTA
      AAAAAGATAGGAACGAACTGAATATAAATTCGTCGTCCTACCTCAAGATTT
      CAGATATTATTTAAATATCTAATTTTAAAGTGAAATAATCTCTTAAATTAA
  15  TGACTTAAATTTAAAATTTAACATAATAATTACAATAGTTAGTAAATCGTG
      TTAAGTAGATAAAGAATATAACTTTTATAATGTTCAATACACTTCATTTTA
      CTCTTTATCATTGTTAATATTTAAAACTCGAGATTAATCATTAAATGTTCG
      TTTTTCAATTTTGCTACTAGTTTTTTCTACTTTTTGTTGTTTATTACGTTT
      GTTTTGGTTAGGTGTATGTTGCTGTTTGCCTTTGGAATCCTCTTTAGTTGC
  20  TGTTATATTCTTAGTCTGATTCGAATTTTTTTGGCTGTGCGGATTATTTGA
      ATGTTTTTTAGATTCAGTAATAAGTGTAGGGATACGCTTCAATTTTGTTTG
      AGCTTTAGCAATTGATGGTTCTTTTAACAACGCTTCTGTAAGTTGTGTTTC
      TTTAGAGCCATTTGTGAGCTTAGATCTTTTTTTCTTTTGCTGCCTGTGCTAA
      GCGATAAGCAACTGAATTAGTATCTGATGCTTGAGATTGATTTAAATTAGT
  25  TACAAGATTTGGAACATCAGAGTTATTTGAAGCTTTTATAGCTTCTTCTTG
      CATTAGATTTTTCGTGTTACTAGCAACTTCATTATTATTATTTATTCTATC
      TTTTTTTGCTTTTGCCGCTGCAGCTAATTTTTCTGTTTCAGTTTCTGACTT
      ACTTTTAGTATCTTCTCCACTTGCAGAAGGGTTTGCTAAATCATGATTGCG
      TGCTTCTTCTTTAACTGCCGTTTGAGCAATTGCAATTTCTTCAGCAGAAGG
  30  TTCTTTCTCACTATCTATGTGCGTCGTAGATGTTTTATTGTTTGAATCAAC
      TTCACTTTTAGCTTCTTGAATTGCCTGCGCTTCCGGTGATTCATCCGCTAA
      GTCTGTTTCGTGTTGAATAGCATTTTGTTGAGCTGCTAATTCTGTTGCTGA
      TGGTCTCACATCATCCACTTGTTTTGTTTCTTTTTTTGCTTCTTGAATTGC
      TTGCGCTTCCGGTGATTCATCCGCTAAGTCTGTTTCGCGTTGAATAGCAGC
  35  TTTTTGAGCTTGTAATTCATCTGCATGGATATTATCTTTGGAATCATATTT
      TAATTGCTCAACTTTAGATTTAACAGAAACCACTGTATCTTCCGCTCGTTG
      TCTATAGTTTTCAACTTTACTACGTAAATCTTGTTCTTTTTCTTTCAATTC
      ATCCGTTTTTTTATATATTTTATTTTTATAGTACAATCCAACAGCAGAACC
      TACTATCGCACTTGCTATGAAACTAACTACAAAATCCTTACGCTTCAATGA
  40  GTTAGATTGTGAACGAGATGTGCGATGATACAATTCTTCGCTTGTATGGCT
      TTTCTCGAAATTATCACGATTATAATGTTCCATAAACTTTCCTCCTAAAAC
      AAAAGCACTTAAATAAGTAAAAGCCTCATTTAAGTGCATTTGTCATTAAAT
      AGTATTAATTATTTATTAACACGTGAATTGTAGCTATGATTTGCATCATCT
      GCTACAGTGTTTGTTTTGTAGTTTGCACTTCCTCTACGATTGTATCTGTTT
  45  TGCCATTTGTCAGCAATTTCCATTGCAACATTTGACCATTGTACAACTTGA
      GAAATTTTATCTTCATTTTGAGAAATATTGTGCGTAATTGAGTTTGTAACT
      CTATCTACAGATCCATTTAAGTTTTGTACTGAGTCTCCAATGCCCTTAACA
      GCATCAACTACAGAATTTAAACGATCAACTTTACCTTGAATATCTTCAGTT
      AAGCGATTCACTTTATGAAGTAAATCAGTCGTTTCACGAGTAATACCTTGA
  50  ACTTGACCTTCAACACCATCTAGTGTTTTAGCCACATAGTCTAAATTCTTT
      TTAACAGAAATAAGTACCACGACGATACCGATACATAAAATTAAAAACGCT
      ATCGCAGCGATGATTCCTGCAATTGGTAAAATCCAATCCATTAAAAAACGCC
      TCCTAATTACCATAAAATAAAAGTCATTAATATTTAATACCCATGATGCTT
      TTATATAAACATATTAAAAATCATTTTTTACGCCTAATTTTTCGAAATAAG
  55  CTTTTAGTAACTTTTGAATGTCACCAGCGCCCATAAATAAAATTACAGCGT
      TATCAAATTTTTCTAATACATCAATACTATTTTCATCTATAAGCGTCGAGC
      CGTCAATACGATTGATTAAATCTTCTATAGTTAAATCTCCCGTATTTTCTC
      TTATTGAACCGAATATTTCACATAAAAAATACTTGGTCT
```

Sequence 3424
step.1002h09.cons.ok
GAGACGTCCAAGTTGAGGATATGTTTCAGTATTACCATGAGTCATACCCAT
5   ACCGCCACCGATAGTCACATTAAATCCAATTAACTCATCTTGTTCAACGAT
AGCGATTAAACCAATATCTTGAGAATAGACGTCAATATCATTAGATGGTGG
TACTGCAATACCTATTTTGAATTTACGTGGTAAATACGTATTCCCATAAAT
AGGTTCCTTTTCCTCACTCGAATCTAAAACCTTTTCACCATCAAGCCAAAT
TTCATGATATGCATTTGTTCTTGGAAGTAAGTGATTACTTATACGCGTTGC
10   ATAATCATTAATCTCCTTATGTACTTGAGATTGATAAGGATTAGGATTGCA
CATCGTATTACGATTAACATCTCCACATGCAGCAATAGAATCAAGTACTGC
ATGATTAATATTTTTCATTGATTGTTTCAAATTACGTTTAAGAATTCCATG
AAATTGAAATGCTTGGCGTGTTGTTAATTTAATCGTATGATTTGCATATTG
ATTAGAGATATCATCCATAGCAATCCACTGTTCAGGAGTCGCTTTCCCCCC
15   AGGTACACGAACTCGAATCATAAAACTATATGCAGGCTCAAGTTTTTGTTT
ACGACGCTCATCTCTTAAATCCCTGTCATCTTGCATATAACTTCCGTGAAA
CTTTAGCAGTTTTGCATCATCTTGTGTAATGGATCCAGTGATTGGATTAGC
TAAACTTTGTTCAATAGTTCCACGCAAAAAGTCACTATTTGCTTTTAAAAA
TTCCATTTCATCAAGATTTTTATCTAATTCTTCCGAAATATGATTATTTGT
20   ATTAACCATGTATCCCCCGTCCCTTCTATTAATACACGTCTCTTTGATATC
TTTTATCTCTTTTCATTTGTTTTAAGTATTCTTCTGCATCTGTTTCAGATA
GGTTTTGCTCTTTGATTAACACATTTTTAATCGCTTGATGAACATCCTTTG
CCATTTTACTTTCATCACCACATACATAAATAGTAGCGCCATTTTCAATCC
ATCGATTAAATTGTTCACTATTTTCTGCAATTTTATGTTGCACATACACTT
25   TTTTATCAGTATCTCTAGAAAAAGCAACATCTAATTTTGATAAAGTTCCAT
CTTCAAGCCATTCTTGCCATTCCGTTTGATACAGAAAATCTGTAGTGAAGT
GTTGATCTCCAAAGAATAACCATGTATTTCCTTCAAAACCTAGTTCCTCTC
GTTCTTGCATATAGGATCTAAACGGTGCAACACCTGTCCCAGGACCTATCA
TAATCACAGGTGTTGATTCATCTTGCGGAAACTTAAAATTCGGATTTCGTT
30   TTAAATAGATAGGAATTGTATCGCCCTCTTGTATTCTCTCTGCAAATTGTA
CTGAACAAACACCTGAACGTTCCCGACCGTGTGCTTGATATCTAACTGCTC
CAACAGTAATGTGAACTTCATCTGGTGTTGCTTTATAACTACTAGATATTG
AGTACTCTCTAGGTGGTAACTTTCTTAATAATTGATGTAAATTTTCAGGTT
GTAGTTCTGTCGTTGCGAAGTCATTTAATAAGTCAATCAAATCCCTTCCCT
35   CAACGTAGTTTTGAATCCATTCTTTATCTTGAATTTTTTCAGAAAGCTCTT
CATTATCAAAAAATATCGCAGCATTTTCTATTAACGGTTTTGTTAATTTAG
TAATTTCAAAATGCGATGTTAATGCCTCTTCAAGATTTAAAGTATCTCCAT
CCTCGTTAATTAAAACTTGTGTTTCTGGACTCCAACCTAATGTGCTAATAA
GTAAGTCTACGATAGCTGGATCATTTTGAGGCAAGACAACTACACAATCTC
40   CCGGTTCATATTCCTCACCAAAATTATCAAGTAATAACTCAATATGTCTCG
TCTCTTTGTCTGAACCTCTACCATTCAAATTAATATTGGTCAATACTTCAG
CATCGTAAGGATTAGACTTACTATATTTCTTTTCTTTTGCTGATTTATAG
ATTCGCTAACAACTTGTTCACTTTCAGTACCAGCCGGTGTTGAATCAATCG
TATTAATAACATTGGCCATCCACTTTTCAGCGTCTTCTTCATAATCAACAT
45   CGCAATCAGTACGATGATAAAGTCTTTCTGCCCCTAATTCAGCCAAACGGT
TATCAAAATCTTTACCAGTTTGACAGAAGAATTCATAGGTTTGATCTCCTA
ATGCTAATACTGAAAATCTCACCCCATCCAATTTTGGCGCTTTACGTCCGT
GGATGTATTCATGTAATTCAACAGCGTTATCTGGTGGATCTCCTTCGCCAT
GTGTAGATGTGATAATAAATAAATCTTCTACTTTCTTCAAATTCTTAGGTT
50   TAAAGTCATCCATTGATTTGAGCGTAACGTCATTTCCAATATCAGATAAAC
GTTGTTCAAATATTTCTGCAAGTCCTTGCGCATTACCAGACTCAGATCCAT
ATAAAATTGTTATCGATCTAGCTTCCGGCTCAACTGAAGGTTCCTTTTCAT
GTAGCATCGCTTCTGTATTGTCATCTGAGTTATGTTGTTCTACAGAATCTG
TAGATGTATTTGATTGTTGATTTGCCATCAGATAGCCACTTAACCATACTT
55   GTTGGTTAGATGATAAAGTGTGAAGCAATTCATTGATTTGTTTAGCTTGCC
CTTCCGTAAAAGGACTATTGGTAACACTTAAGTTCAATCTTATCCACCTCT
TATACTTTGTTATTCCGATATTTGTCATTATTTATATATGTCTATTTTTGA
AAGTTATTAATTATATCTTCAATTCTAGTATGTACCATTTGTTATTTTCAG
TAAATGTAATACTCAATTGATAACTAAAATCGAAAAGACAATTCATAATAT

```
     CAAAGGGCAATACAAACTTTTATTACTCAAAATTATCGTTACTAAATAAAG
     CTGTATCAGCTTTTAACAAACAACTCTTTATTTCTTATCGGAATACTATGT
     TTTGAAATTAAAATTTTTTGACTATCTCTTATGGTTTATCAGCTACATGTA
     ATCCACATTCAGTCTTACTAGAATTGGACCAACGACCAGCACGTGAATCAT
  5  TAGAATCAAATACGGGTGATGTACATGGAATACAACCAATACTTGGATAAT
     TTTGATCATGTAATTCATTATATGGTAAATCCTTATCACGTATA
```

```
     Sequence 3425
 10  step.1002h10.cons.ok
     TAGGTAAACAATGATTGATACTTTGATTTGTTCATTGCTTTATCATTTAAA
     CAAACATAGTGATTTTAATTTTTATCATATAAAAGAATGAAGATAAAAAAG
     ATGTATGTGAAAAGTGTGCTCACATACATCTTTTTATTAATATACATCTTA
     GTTAGCACCAACTGCTACTATAGAATTTGATGTGGATGCATCATCGTAAGA
 15  ATCAATTACTTTGAATTTTTTACCGTTAACTGTAAATTGTTTGTTTTTAGT
     ATTGTGATATACCCAAAAGCCATTTATTTCAGCTATATCCCTATCATTAAT
     TTTTAAATCCGTCATGGTCGAACACCCTATACTCATTAATTTGATTGTTAC
     TTATGTTTGGATCGTTGATTTTTAATTTGTTATTGCTAAATAATAAGTGGA
     ATGTGAAATTATGTGGAATAGATGGCTCTTTTCTCATTTCTTTAGCCATGG
 20  ATAAAGTATTTTTTCTAATGATTTTTTCGTCATATTCGTCCAGAGTACTGA
     AAAGTGTAGTCGTTATTTCTGTATTAATATACGTATCTTTTTTAGACTTTA
     ATTGTTTCATATGCTTTTTAAAATCATCGTCACTTTCTTTAATTAAAGGTT
     CAAAATGTTTTCTATAATCTCGTAATTTCGTGGAACTAGTAGTAATGTAAA
     AGTAATTATTTTGATAACCTATATTTTGAGTTTTAGTTATTGCTTCATCGG
 25  TGAATCCAGTATAGTGATATTTATCTTCTTTATCTTTGAGGTATTTTTCTA
     AGTTATCCATTTTTTTCCTTATTTGCTCTATACTCAAAACCGCTTAACACAG
     TTCCTACCATCATACTCATAGTATCTCCGTTATCATTGCTTCGCATGGAGC
     CTTCTTCATGGATTGATTCCATATCAAAAGGGATGCTTGCATTAAACACAA
     TGTCATGATCATCGCAATGAACGTAAACTTCGGCGCCATCACCACTTCCTA
 30  CAACATTTGTGGCTTTAACATTAAGGCTAAAATTATCTTTAAAAAAATTGTT
     CGCCAAGTTTTTCGTATTCTTCACGATATTTTTTCGCATGTGCGATAGCAC
     TTTTTTCAGCTAATGGTTGGAAATCTTGACCTTTGTATTGACTTGCAGGAA
     CTTTTTCTATTTCGTCTTTCATTGTCGAACAACTTCCTAATATTAGTGTTG
     AAACGATAATTGTGGCGATTAATTTTATCATGATACACCTTTCTTTACTAA
 35  TATTACATCTTAATTAATATGAAATACATTATTTAAATCAAGCCCAAAAGC
     TATTTAATTCAGCAATATCTCTATCATTGATTTTACTCATGTTTATGCACT
     CCATATTCTAATGGGTTGTTAGCATCATAAAAATGCGCGTGTAGAATTAATT
     TTTGGCTCACTTAATTGTAAAGTTATGTTTGTGTTTTCAGGTTTTTCTTTG
     TAATGATATAAATCATCACTTAAGAATAATAAATCTTTCTTAGAGTTATTT
 40  TTAGTGAAATCAGATTTAGTACTGAATAATGTTGAAACAACTTCAATTTTG
     GCTTTGAAATTTGTTGCTTCATTTGCTTTTTTTCATGCCTTCTTTGAAATTC
     TTAGAATCTTTTTTTTATTAAAGGTTCATAATATTTTCTATAATTATTAATA
     TTAGTAACGTTACCTGCTAAATAATAATACTCATTTTGATATCCAACATTC
     TGTGTCTTATGTATTGCTTCTTTAGTAAAACCGGTATATTGATATTTTTCT
 45  TCATTGTCTTTAAAGAACTTTGTTAAGTTATCATATTTTTTCTTTTTGTGCT
     TTATATTCAAAACCACTTAACACCGTTCCAACCATATTACTCATGGTATCT
     CCGTTATCATTGCTTCGCATGGAGCCTTCTTCATGTATTGCTTCCTTATCA
     AAAGGGATGCTTGCATTAAACACAATGTCATGATCATCGCAATGAACGTAG
     ACTTCGACGCCATCACCACTTCCTACAACATTTGTGGCTTTAACATTAAGG
 50  CTAAAATTATCTTTAAAAAAATTGTTCGCCAAGTTTTTCGTATTCTTCACGA
     TATTTTTTCGCATGTGCGATAGCACTTTTTTTCAGCTACTGGTTGGAAACCT
     TGACCTTTGTATTGACTTGCAGGAACTTCTTCAGGAATTGATTGTGATGAA
     TGTTTTGCTTCTTTATGCACTGTGGAACATCCTCCAATTAATAATGTTGAA
     ACGATAATGGTAGATATCAGTTTTCTCATGATACACCTCAATATAAGATTA
 55  GTTACTTTAATTTTAACTTAATTTATAAATTATATAAAAAGAATTTATAAA
     AGTATTGAACAAATATAACATTTTTCACATATTAGTTATATTGATTTTATA
     GCAATTTCAAACATATTCGAATAAAAAAATAAAAAGCTTACACAATGTAAT
     TTGATACATGCATAAGCTTTGTCATATGATTAACGATTTTTATTCTGTTTA
     TCATTTGTTCCATGATTTTGATTGTTGGAATTATTATTTGAGTTGCCATTT
```

TGTTGTGAATCAGATTGATTATTATGTTGTGACGACTGTTGTGGTGTTGAT
TGTGCTGGTTGATGATATGATTGATTAGTTTTGTTCTGATTCGGTTTCTGT
TTGTTATTTTCTGAAGAGTTATTTTTATTTGAACCAGACTGTTGATTATTA
TTTGGGTATTGATGGTAGTTAGGTTGTTGTGCTGGTTGTTCAGTACCTGAT
TGATTATCGTTGTTATAGTCATTATTTGTTGGTTGAGTTTGGTCTTGGTGT
TGATTAGTATCACGATTTTGATTTGAAGTATCATTTTGTTTGTTTTTGCGA
TTGTGTTCTATTTTTTTTATCATTCTTATTAGAATTAGTATGGTGGTTAAT
AGCTGCATATAATGAAAAGCCTAATGCACATATAAGAATGATTAC

Sequence 3426
step.1002h11.cons.ok
TAACCCAGCCACTCGAACCATAAAGAATTCATCTAAATTTGAACTGAAGAT
AGAAATAAAGTTAAGTTTTTCAAGCAACGGATTATTTTTATCATATGATTC
TTGTAATACTCTGTAGTTAAAATCTAGCCAACTTAACTCCCTATTGTTGTA
ATACTGCGGTAAATTAATATCTTTTTCTCCCAATCGAGTTTGCATACTATA
CATACACCTCATTAATTAATCTTTATTATGCCCAATCTAAAATATCATAAT
AATTTTAAGATTTTGTAAATATAATAGAAACCTTACCTTTTAAAATTTTCT
CAATATGCTTTTTCTGACGATTTGCTTGGTATTCTTCTGCAATAGGTGAAC
CTTTGTAATAAACTAATAAATCGTATTTGTCATCTTTCTTTGCTTTTAGTT
TAACTTCCTCTACAAAACTAGTATGTGAGATATTCAAGGTGTTTGCAAATT
TAATAATCCCTCCTAAAGCTTGTATTGTATCTATTTCTTTATTACTAAACC
ACTGTGTTTCTTTGCAATAAAATTTAAGTAAAGATTTGTTTTTAAAACTAG
CTAACAAAGCTAATTTCACACGATCTTTATGTGAAAAGCCGTTAATCATTG
AATTTGCGATTAAGTAATACGTATGTGGTGAACTTGAGTCTGAATCAATGA
AACTACCTAGATAATAAATGTAGGCACCTTCAATAAATAATTCTTTTTCCA
TTTCTGAAATATTAAGTGATCGTTCACTTATAATTTGATTCAATAAGGATT
GAGCTAATTTCACACGACGATTAGCACTCGTTTCTTCAATATGATATTCGT
TCGCTAAATGACGTAATGCATCTTTACGTACGTTACTTTTATTAAACTCAT
CAGGATATCGTTTGCTGATGTGGTTCATAATAAATCCTTCACGAATTCCTT
TTCTTGAGAAGGTGAATTGTGTGGCGTCAATTTTTTTGAAAAGTGTTTTAA
AGACGGAGATAGCCGGCAGAATAATATCGACGCGATCGCGACTTAAACCGT
CTAAATTTGTTAGTTCATCGCGAGAACTTTTACGGATTAGATCATAAACAT
TGTTAATATCTTTCGAAGTCATCGTATAGTTATGAACGCCACCGATAGGGT
ATGCATGTGCTGATTGATGAATGCGTGCAACATTACGTGCAGAACCACCTA
CTCCTACGAGCGCAATATGTTGGTTGGATAGCCAGTCTAACTGACTAAATT
GCTCACGTAAAAACTGTTCCATATTTTTAATGGCTGTTTTGTCATTATGTT
CTTTATCACCAAAAAACTGACGCTTAAGTGATACCACGCCAAATGGAAAGC
TATGAGCCTCTTTAAGTTGTTTGTCTTTGAAAAGGGTAACTTCGGTAGAAC
CGCCACCGATATCGACAGAAATTCCATTTTCAATATCAGTAGTATGTGTAA
TCGCATAGTAACCGTAAAATGCTTCATCTTCTTCAGGTACAATTTGAATTT
CGATATGAATATCTTGTTTAATTTCTTTAATGATAGCTTCACGATTTTTAG
ATTGACGGATAGCAGCTGTTGCGATGGGATATAATGCATCAACGTTAAATT
TATCCGCAACTTTTCTAAAACTGCTTAAAGTCTCTTTTAAAACATGAATAC
CTTCATCATTCATTTCATTGGACTTAGTGAGATATTGACTTAAACGTGCAG
GTGTTTTTATATTCAGTATTTCATTGAGCCCAGTTTTTTTTATTGTAGCCAA
ATATAACAAGTCGAATCGTGTTGGAACCAATGTCTATCAAACCAATTCTTT
CTTCCATGTTTGCCTCCTATTATTCAAACAATGCGTTTAAAATAATCCCTA
TTTAATTTTATTGTACATGATATAGATAAATTGAAAAGAACTTACAGTTAA
TTTTTCTGAAATATTTAAGATGATTGATAAGTAATTATAATTTTTAAAAAT
ACAATTTTAGTAAAAAGAGAAGTATATAAATGTATAGCTTAACTCAGCTAT
TCATTCTTCGTGCGTCTTATTCTTATCACTGTTTAAGTTGATTCACATCGT
GGAAAACTAAAACTAACATGATACCTAATTGTATTAAGTATGTACAAGTGA
TAGGAGAGTGACTTTGGATGGAACGTTTAGAAAACAAAATCGCAGTGATTA
CTGGTGCGAGTACTGGTATTGGACAAGCATCGGCCGTGGCGTTAGCAAAAG
AAGGAGCACATGTGTTAGCGCTTGATATATCAGATCAATTAGAAGAAACTG
TGCAGTCTATTAATGATAATGGTGGGAAAGCAACTGCATATCGCGTAGACA
TTTCAGATGATAAACAAGTCAAACAATTCTCAGAAAAAAATAGCACAAGAAT
TTGGACATGTAGATGTTATTTTTAACAATGCGGGTGTAGATAATGGCGCCG

GACGTATTCATGAATATCCAGTTGAAGTGTTTGATAAAATTATGGCTGTTG
ATATGAGAGGAACTTTTTTAGTAACTAAATTTTTATTACCTTTAATGATGA
AACAAGGTGGTTCTATTATTAATACAGCTTCATTCTCTGGGCAAGCTGCGG
ATTTATACCGTTCAGGGTATAATGCTGCTAAGGGCGGTGTCATTAATTTTA

5   CAAAATCTATCGCTATAGAATATGGACGTGAAAATATTCGTGCTAATGCTA
TAGCACCTGGAACAATCGAAACACCACTTGTTGATAATTTAGCAGGTACAT
CAGATGAAGAAGCCGGACAAACATTCCGAGAAAATCAAAAATGGGTAACAC
CATTAGGTCGACTAGGAACACCGGATGAAGTTGGGAAACTTGTAGCCTTTT
TAGCTTCCGATGATAGTTCATTTATAACTGGTGAAACTATTCGTATTGATG

10   GTGGCGTGATGGCTTATACATGGCTACAACACGCATTTTCATTTTTGTCG
TTGTTTTTTCTTATTTTCTTGTGACGTATATTGGTTACCGATATGTTCTAC
TTTTTTATTCATTGCTTGTCACCTCCTTAAGCATTTCACTCTTCATTAATA
CGTTCTTCTTTAATC

15

Sequence 3427
step.1002h12.cons.ok
TGTTCTAAATTAGCAATAACACTTTCAAGAGACGCACGTGTAGGATTTGCA
GTACGTGAATATTCGTACCCTTGTCTTAAATCACCAATATCATCTTGTAAA

20   TAAGTACTTGTTTGATAAATAGGTGTTGTCACTGCTCCAGTATAGTTGTCT
GTCGTATGTCCCCCATGTATCATTTGCGTTTTTTTATTCATTATTAAAACT
CTCCTTATAGTTGAATATTTGTTTGGACATGTATCGATCGCTTCCATCTGG
AAAGATGGTAACAATCACACCATTTTGAATGCTTTTTTTTAATTCCAACGC
TCCTTGTAATGCCGCTCCCGAAGAACTACCAACTAACAATCCTTCTTTATT

25   CGCGACAAGTTTAACATTATTAAAAGCATCTTTATCGGCAACAGTAAAAAT
ACCATCTACTAATTCTTTTTCTAAAAATGATGGCCACTTTTCAGAACCAAT
CCCTTCAGTTGCATGAGGATGACTGACACCACCATTTAAGACAGAGCCTTC
TGGTTCTACAATATAATTTTTTACATCATACGTTTTTAAGTGTTGTGCAAC
TCCTGTAAACGTACCACCGGACCCAACACCTGCCACAAAATAATCAATATG

30   TGAAAGTTCATCTGTGAGTTGTTTGGCAAGTGTTTGTGTATATGCCCCAGG
ATTATCTTTAGTTTCGAATTGATTCATATATAAATATCCATATCGTGTTGC
GTATGCCAACGCCTCTTGCTGTGCGCCAGTCATTCCCTCAGCTTTGGGGGT
ACGTCTAACATCTGCACCCAATGCTTTCATAATTGAAATTTTTTCTTCTGC
AAATCCTTCTGGAGCAAAGATGATACATTTTACTTTGTGCCGATTAGAAGC

35   AATAGCAAGTCCAATGCCTGTATTACCAGCAGTCGCTTCGACTATTGTATC
CCCTTCTTTAAGTCGTCCTTCATCTATTGCTTTTTTCAATTAAGTACTTCCC
TAGACGGTCTTTGATGCTACCACCAGGATTAAATTGCTCAAGTTTGGCGTA
TATTTTAACATTCTCGTCACTAAAGCTTTCTAATAAAACTAATGGAGTTTG
TCCTATCAAATCGTATGCAATCATTATTTTTGGTTCAGATGACAACAATGA

40   GTATCACTGAACTTAACCCACTTTCTTTATTTATGTTAGCGCTAATTCTTA
CTTATCAGATATGAATTATAACTTATCGGCATTGTTTTTGCAATTTCTGAA
AAACGTACAATCAATTTGAAATTTAAACATTATTTAAAATAAGTATTTTTA
TAATTGATTACGTTTATATGAGAATACATACCGAAAGTGTTATGATAGGGG
TAAATAAGCTAGCACATCGATAAAAGTTATCGAATTACATACACTAGAAAT

45   GATAAGGAGCTGCTTTAATTATGACGAAACAAGATTTATCTTTATCTGTAT
TTACCAATGAAAATTATAAGAATCTTCGTTACACATCATCTAGTTTTAGAA
ATTCTATGTATGATGAATTAGAGGTTAATAAAAGTCGTTTTAAAAACTGTA
ATTTTAATGAAGGTATTTTTAAGAATATAGAAGCAATTTGTAATTGTAAAT
TTACAACGTGCGGGTTTAATAATTGTATTTTCGAAGATGTTCATTTTTACA

50   AAAACCAATTTAAAGATTCAACATTTGTGAATACACCATTTGATCAATCCG
TATTTAATAGTACTTTATTCCAAAATGCAATGTTCGATAGCAATCTCATTC
GTAGCGTAAAATGGACTGATATCATTTTTAAAAACGTTTCTTTCAAAAATG
TAGAAATTGAAGGAACAACATTTAAAGATGTAAAATTCAAAAATTGTGAGT
TCAAAAATGTAATTATTACTAATTCAACTATGTCGCAAAAGTTAATGAATG

55   AATTACAAAAACAAGATGTTACTTTAGAAAATATAGACACTTCTATTTAAC
ATTCATTATCTTCCGCTTAAGAAAACAATTTAAGATCATAGTTAAACATTC
GCTCGTTTAATGGTGTTAGATATTTCGAATCAACTACATAACATTCTTTAT
ATAGTAAAACGGTCAACTCCACTAAAAAATGGAATTGACCGTTTGTTTTAG
TTTTATCAATAACGTCTCAATGACTTAAAATAATTGAACTATAAATACTAA

```
AATAATAACAATAGGCATGATGAACTTAATTAAATAATACCACGGTAAAAA
TAGGTTGAATTTGTCTTTCCCAAAGCTTTCTTTTAATAATTTTTTATCTAG
TAATTGGCCAACCACTAATGTTGTTCCTAGTGCCCCTAATGGCATAAGAAT
ATTAGATACAATAAAATCCATATTATCAAATATCGTACCAGCGCCAAAACG
CAAATGACTTAGACTACTAAAAGATAATGTTGCTGGGATACTAATGATAAA
TACAAGTATACTACCTATGATTGCCACTTTTTGTCTTTTACTATTATCATT
TTTAGTAAAATTAGATACATTTAACTCTAATAATGATATAGAAGACGTTAA
TGCCGCAAATAAGAATAATAGTAAAAATATAAAGTAAAAGAATGTACCAAA
AGTCATTTCGCTAAATACTAGTGGTAAAACCTTAAATAATAAGCCAGGGCC
TTCTTGGGGTTGGTAACCAAATGTTTTAAGCGCAGGAAATATAGCTAATCC
AGCCAAGACAGAAATTAAAATATTCATTACGACAATTGAAAGTGCTGAAGA
CTTTATCGTCATATTTTTAGGTGCATAGCTTGCATAAGTAATCATTCCGGT
TGTACCTAGGGACAGCGTAAAAAACGATTGTCCTAACGCAAATAGTACACC
TTGAATAGACATATCTTCAACTCGAGGTTGCAGTATATAACGTACACCTTC
TAAAGCACCTTCTAAAGTTAAAGATTGTGCTACAACGATAATTAAAAAGAT
AAATAATAGAGGCATCATTATTTTAGAAGCTTTTTCTAAACCTTTTTCAAC
ACCTAACATAACAATTACCATTGTTATCAAAATAAATATGCCTTGACCTAA
AACAGTCAACCATGGATTACTAATGATTGTTTCAAATTGAATATTTGTCAG
CGTACTTGATTTAAAAACCATGATTTGTGCTATGACATAACCTATATA


Sequence 3428
step.1003a02.cons.ok
AAAATCAAATCCGGTTTAGCTTTTATTAATTCTTCTTTATTTAAATTCATC
GCATCGAATTGTTTTTTACCTTTTTTTACATCTTTAGGATAATCATCCACA
GTGGATACACCAACTATATCTTCTCCGATTCCTAAGCGATATAAAATTTCT
GTGTTACTAGGAATGAGCGAAATAATTCTATGATATTTTTTATCGCTTTTT
TCAGTGTCTTTATTATTAGATGAATGATTGGCATCTTGACCACATGCAGCC
AGTACTACAATCAATGCACATATAACTATCCATAGTTTACTTTTCATTTGG
AGCATCACCCTTTCACAAATTATTTTTAATGTTCCTATTCATCATATCAAA
ATATACCAATATTTGTAGATACAATTTTATTTGATTTTCACAAAATATGAA
AAAAGACAACTTTGAATTAAATCTATTAATATATTTAACCCCAACTTGCTT
TGTCTGTAGAATTTCTCGTTGAAATTCTCTGTGTTGGGGCCACAACCCCAA
CTTGCTTTGTCCGTAGAATTTCTCATTGAAATTCTCTGTGTTGGGGCCACA
ATCCCAACTTGCTTTGTCCGTAGAATTTCTCGTTGAAATTTTCTGTGTTGG
GGCCCCGACCCCAACTTGCCTTGTCTGTGGAATTTCTCGTTGAAATTTTCT
GTGTTGATCAGTTTTATAATGTTAGTATCAATCGGTAAAATGCTAAAATGT
CTGAGACAATCAACAATGTCCCAGACACATAATGTTTTATAATTTAATAAA
TGTTTTTATCGCGTCAGAAATGTATAAATTCAAACCACGTTCACGTTGCTT
GTTAATAAACTTTACAAATCGCTCGTCCTCAATATAAATCTGAGCTATATA
TTCTAAAAATTGATTGTCACAATTAGGAACTTGTTCTTTCAAAATACATTG
TAATTTCTTAGTTTTTCCACTTGCTTCAAGTATAGAAACTTTGTTTAAGTA
TAGTTGATTCATTTCGTCAAAAAACATATTTAATTGCTTATTAATTTCCTC
AAAATGATTTGCTGTTCTGATTCATCCTTACATTTTTGTTTATCTTTATA
TGCTTGATAATAATGAGTATCTCCATACTTACTTGCTGCTTCTTTTTCATA
CTGTTCATTTAAATTGAAATTTTGCATTGTAGTTACCTCAATCTCATCGTT
ATCTTTGTGAAAATGATGTTCTAACGTACGGATAATATCGTTTAAGCGATC
TCTTTTTTTAACAATGTGATAAAAATTATCATGTAATATTTTAGTTCTAAA
ATCACTATCACCTTCAAAATACGTCTGTATTTGCTTTAATGTTAAATCTAA
TTCCTTTAAAAAGAGAATATATTGTAACTTTGAAATATCTTGAACATTATA
ATATCGGTATCCATTAGCACTCATTGTTGTGGCACGAGTAACCCCTTTTC
ATGATAATAATGAAGCGTGCGTATACTAACTCCGGTAATGTCAGAAAGCGT
TTTGGGACTGAAGCGTTCTTTCATCATTTCACCTCCTGAATTAAAGTATGA
GGTATGACGTTACGTTAAGGTCAATACTTATTTTTTAAAATAGGTATCTAA
AAAGGTATAGCGTTGTTTCCCCTTACCATTAAAATAATTTTTCCCAAAATT
ATCTAATAAAGCGTGAAACTCATTTGCATCTTGATTTGAAAATGATTCAGG
AAGTGTTAATTCCTGTTTCAATTTATGATAACTTTCTGTATGTTGATATCC
CAATTTTCTAAAAATACGTCTAGTATAACTATCAGGTATGAATTCTTTACC
TTTAAAAATATATACTATTAAGACATCGGCGGTCTCTTCACCTATACCACG
```

```
GATGGTGAGTAATTCTTTTCTTAAGCTATCACCGTATAACTTAGCTATACT
ACTATAATCAAAATGATGTTGATTTAACCATAGTAACAATGCCTGTATAGC
CTTAGCTTTATTTTTATAGAAACCACTCGAACGTATCACTTGCTGCAACGA
TTCTAAAGGCATTTTCAATATCGTCTGTGCATTAAAAGAAGTTTCTTCTTT
```
5
```
TAATCTTGATAACGCTATATCTGCATTGTTCCAATTAGTATTTTGGACTAG
AATTGCCCCAAGCATCATTTCTATTGGCGTTTCTGCTGGCCACCAATATTG
AGGGCCCATATTGTATTTTGAGGAACAATAGGTATCTATAATGATTGATTT
AACAATGTTTGTAACATGTTAACAGATTTTACAAAGACGGTAGCTAAATAT
ATAATATGGTACATAGATGATGAAAGTGAGATGACATTATGGGACGTAAAT
```
10
```
GGAACAACATTAAAGAGAAAAAAGCCCAAAAAGATAAAAATACTAGTAGAA
TATATGCCAAATTTGGTAAAGAAATATATGTAGCTGCAAAGTCTGGTGAGC
CTAATCCAGAGTCAAATCAAACTTTAAGATTAGTATTAGAACGTGCAAAAA
CATATTCAGTACCTAATCATATTATAGATAGAGCTATTGATAAGGCTAAAG
GCGCTGGTGACGAAAACTACGATCACTTAAGATATGAAGGTTTTGGTCCGA
```
15
```
ATGGTTCAATGCTTATAGTTGACGCATTAACAAACAATGTAAATCGTACAG
CATCAGATGTACGTGCTGCGTTCGGTAAGAATGGAGGAAATATGGGAGTAT
CTGGTTCAGTAGCTTATATGTTTGACCATACTGCAACCTTTGGTGTAGAAG
GTAAATCTGTAGATGAAGTCTTAGAAACACTAATGGAGCAAGATATTGATG
TAAGAGATGTAATTGATGACAATGGCTTGACTATTGTTTACGCAGAACCAG
```
20
```
ATCAATTTGCACAAGTTCAGGATGCATTACGTGAAGCTGGCGTTGAGGAAT
TTAAAGTAGCAGAGTTTGAAATGTTACCTCAAACTGATATTGAGTTGTCTG
AAGAGGATCAAGCTATTTTTGAAAAATTAATCGATGCACTTGAAGACTTGG
AAGATGTTCAAAATGTTTTCCATAATGTAGATTTAAAATAATGATGACTGC
TAATGATTGGATAGACCGGTTGGAATTAATTTCGCATCCTGAAAGTTTCAA
```
25
```
AACGATTTCGATCAAACTGTTTATTTGTGTCAAGAAATGTTCCATCCATAT
CTACCGCAATGGCTTTAATCATCGTTAATCCCCCTATCTATAGATTAAAAT
TCATAATCATTATATAGCTGTTGGATTTGGTATTGAAGCAATATCAAATAT
AGTGAAAACATATTCATATATTGAAATTGTTTTCAACATAAAAAATAATTA
CCTTTGTCATTCACACTATCGCTTAATATATAAGATAAATCTCTTATTTTC
```
30
```
ATTTATATAACTTGAATCAAAAATATTAAAAAGAGCCAAGGACATTCACTT
TGTCCTCGACCCTCTAATCAATAAGTTATTAAAATTTATTTTAACAAAGCA
TCTTCTCGAAGAAGAACATAATTAATACAGTTTATTACGTATCTTTCACTT
ATAAGAATATTAGTTAATTTGTGTAGCTATGAATATTTTCCTTGTTTCCCA
AACACCCGTTTCTTCGTTATAGTCATCGCAAGTAATTAATGTTAATTGATT
```
35
```
TTTCTTATTAGGATGTTCGTCTAATACCTTAACCTCTGTAGGCTTAACATC .
ACGTATTTTAGTTATTTTATACTTTCTAGTTTGATTTCCAGTTTTAAAATA
CACTTTACTACCGATTTTGGCTGATTTTAAATTTGTAAATTGATAGTGCGA
ACGATCTGTAAACGTATGACCAGCAATTGAAATATTCTGTTGATTAAGAGA
TTCGTCACCTTCTGCAAAACTAACACCTCTATTGAGTTGTTCTGGTGTTGC
```
40
```
TGGACCAGGGTATACTGGTTCTTTTATTTGTGCATCTGGAACTTCTATATA
ACCAGCCATTTTAGATTTATCGGAAGGTATCTTTGGCGTCGATTTAGATGT
CTTTGTCTGTTCTTTTTCCTTTTTATCATAATTTTCAATTTTATGATCGTT
ATCTTTTTCATGTAGATAATTATCGATATATGGCTTTGA
```

45

Sequence 3429
step.1003a04.cons.ok
```
ATAGTGTTTGATCTAGTGCACCTGCGGATCCAGTATTAATGACGACTTCTG
GATTAAATTTTTCTATTAACAAAGTCGTTGAGATAGAAGCATTAACTTTAC
```
50
```
CTATACCACTTTGTGTTAAAACCACCTCTTTGTGGTTTAGCTTGCCAACAT
AAAATTTAACATGCGCAATATTTATTTCATTCATATCATTCAATTTACGCT
TTAAAATCGTCACTTCTTCTTCCATTGCTCCAATAATTCCTATCATATATC
TTTTTACACCTCATTTATAAAGCATTAGCATAAGTATTTTATCACATTTTA
AATATCTCCACGATAGCTAGTTTTAAAGTATCAATGCTTAAATCACTATGA
```
55
```
TTAAATAATATCTCTTCTTATTTATAATTAAAAACTACCCTGCAATTGATT
TAAAGACAGTTGCGGGTAGTCATAGAACTTTATTATACAATTAACCTTTCC
AAAGACTTGCTAAATCTTGTAATGAAAAGATACCAGTCATTATTGTGACAA
TTACGGCTATAATACCAAAAACAAGCATCCATGTTGGGTGTTGGTAATTAC
CAACGATAGATTTTTTCCTACTTGCAATGAGAATTGCACCTAATGTGATTG
```

```
          GTAGAATCCAACCATTAATCGCACCAGCTATTATAAGTAAACTCACCGGTT
          TACCAATAAATAAGAAAACAAAAGTTGAAATTACAATAAATGTAATAACGA
          TAAGATTATTTTTATTGAGTAACGATTTGTGTAGTGTTTTTAAAAATGTTG
          CGCTTGTATATGCAGAACCAATTACTGAGGACATTGCTGCTGCAAATATTA
     5    CTACGCCAAAAATATTTTTACCTATAGGACCTAATGCATGTTGGAAAACTG
          ATGCTGGTGGATTTTCTGAACTAAGCGTAACGCCAGTTACAACAACACCTA
          GTACAGCTAAAAACAATAAGGTGCGCATGACACCAGTTGTTAAAATACCTG
          CTACAGCAGATCGATTTACGAAAGGAAGGTATGACTTACCTTTTATACCAG
          AATCTAGAATTCTATGTGCACCTGCAAAAGTAATATAACCCCCTACTGTAC
     10   CACCAACTAATGTAATTATAGGTAATATAAGTTTGAAAGGATGTTCAGGTG
          CAAATGTATGTACTAATGCATCTCCATAAGGGGGATTTGAAACAACCATGA
          CATAAGCGACTACTAAAATCATTACGATACCTAGAATCATACTAATAACAT
          CCATTATTTTCTGACCACTTCTACTAACAAAGATAAGTATCGCAAAAATAG
          CTGTTATTGCAGCACCCCATTTTACATCAAGACCAAACATTGCATTTAAAC
     15   CTAAACCTGCACCAGCAATATTACCTATGTTAAATGCGAGACCACCAAATG
          CAATTAGTATGGAGATAATAGTACCAAGTCCAGGTAATACTTTATTAGATA
          TTTCTTGTCCACGTAATCCAGTTACCACTAATATTCTCCAAATATTTATTT
          GAGCGCCTATATCTATAATAATAGAAATTAATATTGCAAATGCAAAACTAG
          CATAAAATTGTGCAGTAAACACTGCAGTTTGAGTCAGAAATGCTGGACCAA
     20   TAGCTGATGTAGCCATTAAAAAAACTGAACCATATAATAATCGTCTATGTT
          TTTTTGTGAATTTAAAATTTTGTCCTTTTTGATTGAAATCTTGTTTTGTAT
          TTTCCCCCATCTTTTGACCCCCTATAAGGATTGAATATCAATGCCTTCTTT
          CATTAATTCATTTCTAATTTGCGTAACAAATTCTAATGCGTGTTTTCCATC
          TCCGTGAACACAAATTGTATCAGCTTTCAAATCGATGATTTTTCCATTTTT
     25   ACTTACAACTTTATTTTCCAAAACCATTTTTAATGCTTGTTGGATTGCTTC
          GTCAGTATTAGTGATAGTGGCATCGGTTTTTTTTCTACTTACCAATTGTCC
          GTCATCTTCATAACGACGGTCAGCAAATACTTCCGAAGCTACCTTAAGCCC
          CACTAATTCAGCTTCCGAAATAAGTAATGTATTCGCTAAGCCGACGAAAAT
          TAGATTTGAGTCGAAATCAAAAACTGCTTGAGCAATTGCATGTGCAATTTC
     30   TTTATTTCTAGCCCCCATTTGATAAAGGGCACCATGAGGTTTAACATGCAT
          CATTTTTACATGATTGATTTTACAAAAGCCACTTAACGCACCTAATTGATA
          TATAACTAAGTTATAAATTTCATTTGGAGTAAGATCCATTTTACGACGACC
          AAAGCCTTTCAAATCAGGAAGACCCGGATGTGCTCCAATACTTATATTATT
          TTTTTTCGCTAATTGAACAGTTTCATTCATAACATCTTCATCACCTGCATG
     35   ATAACCACAAGCAATATTAGCAGAAGTTATTAATGGTAAAATGTGCTGATC
          TCCACCGAATGTATAGTTACCAAAAGCTTCTCCTAAATCACAATTTAAATC
          AATTTTCAACTTCATTTCACCTCTTTAATGATTTGATGTTTTTCAAGAAAT
          TTTATATCTACTTTACTGGCATCGCCTTTTGCATATATTGGATAATTTAAA
          ACCGCATATAAAAAGTCAGCTGTAGTAGTAAAACCTTCGATGACCATTTCA
     40   TCTAACGCAACTTTTAATTTATTAATAGCAGTTTGTCTATTAGAATCCTTT
          ACAATCACTTTAGCTACAAGTGAATCATAATAAGGAGAAACCTGATAACCT
          GTGTAAAGTAAAGAATCTACACGTATATTAAATCCTTGTGGTAAGTGTAAT
          TTATTAACTTTTCCTGGAGAGAATTAAATGCTAAACCAGTTCAAGCTGGTG
          CTTACGATATTCATTTCGTAGACAATGGATACCAATACAACTTCACTTCAA
     45   ATGGTTCTGAATGGTCATGGAGCTACGCTGTAGCTGGTTCAGATGCTGATT
          ACACAGAATCATCATCAAACCAAGAAGTAAGTGCAAATACACAATCTAGTA
          ACACAAATGTACAAGCTGTTTCAGCTCCAACTTCTTCAGAAAGTCGTAGCT
          ACAGCACATCAACTACTTCATACTCAGCACCAAGCCATAACTACAGCTCTC
          ACAGTAGTTCAGTAAGATTATCAAATGGTAATACTGCTGGTTCTGTAGGTT
     50   CATATGCTGCTGCTCAAATGGCTGCACGTACTGGTGTATCTGCTTCAACAT
          GGGAACACATCATTGCTAGAGAATCAAATGGTCAATTACATGCACGTAATG
          CTTCAGGTGCTGCTGGATTATTCCAAACTATGCCAGGTTGGGGGTTCAACTG
          GTTCAGTAAATGATCAAATCAATGCCGCTTATAAAGCATATAAAGCACAAG
          GTTTATCTGCTTGGGGTATGTAATCCAAATCTAATAATATAATTAGAAAGA
     55   GTGAGACAGCATTCGTTGTTCTCTCAGCAACGAGGACTAAAATTGACTCCT
          TGTATATAAGTAAGTTTTAATTTTAGTTATCGATGCAATACATTAAAAAGC
          CTAGGACACATTGCTTTGTGTTCTAGGCTCTTTTTTATTTGAAAGTCTAAT
          TAAAATGATTTATATAACAAAGTGATGTTAACAAAGTGAAAATATAACATC
          TAGACTGAAAAAAATATAATAAAATTAAATGATATTATGTTTGTATTTATTT
```

AAAAGATATGGTTATATAAATGGTGTGTTTAATATCATATATAGATACATT
TAAATTGCTTGCCTACACAACTTATATGAAGATGCTCTAATAAATGTATAC
CTATTTAATTTGTTAAAGTATAATATTAAAGTAAAAAGATGACATGTCAAT
GCAATTGAGGATGTAATTGTGATAATATTAAGA

5

Sequence 3430
step.1003a06.cons.ok
GTATTATATTCAATATTGACGACATAGTCATCTTGATCTTCTTCAATAATA

10   AATGACAAATCAAATTTAGCTGTTGTTGACTGAGGTGGAATATGTGTCAAT
TGACTATGTCCAAAATTCGCATGATTTGTTTCATTATTTTGAAGTACGAGC
ATCACATCAAATAACGGATTATGTGAAGCATCTCTTTCATCAACAAGATCA
TTGACTAAGCTTTCAAAAGGATATTCTTGATGTTCATATGCCCCTAGACAC
ATCTCTTTCATCTCAGCCATCAATTGATCCCATGTCTTTTGATCATGTGGT

15   CGACCACGATATACAAGTGTATTAGCAAACATACCTAACATATTTTCAGTA
TCGCGATGAGTACGCGCACTGATTACACTACCAATAGCGATATCGTCCTGA
CGTGTATATTTGTGCAATAATACCATGATTGCACTAGCAAAGAACATAAAG
TCTGTCACTTGATGTTGTTCTACATAAGATTTCAATTGCTGTTTGATTTGA
CGATTGTAATGAAACGTCAACATATTACCGTTGGTTGTTTTAATACTTGGT

20   CTAGGATAATCCGTAGGCATATTTAATATTGGAACCTGATTTTCAAATTGC
TGTAACCAAAAGTGACGTTGTTTAGATAAGTCTCTGTGCACCATCCACTCA
CTATAATCTTTATACTGAAGCTTAAGTTCAGGTAATGATTTATCTTGGTAT
AAAGCGTTCAAATCAGATAGTAAAATCGTGTTACTCATACCATCATTAATA
CTATGATGAGTATCCATAAATAAATAATCTTGTTGTGGTCCATGTATATAT

25   TTAACTCGCATCTGACTTGGTTGTTCTAAATCAAACGGTTCCATAAATGAT
TGAATAATATCTTGCTCGTTCGTTAGAGATGTCGTTACCTCTTCAAAAATCA
GGCGAAACATGTGTCGCAATACGTTGTTTAACTTCATTGTCATCAATTACA
TATTGTGTTCGTAATATTTCATGACGTTCAATCAACTTAGATAATGCACGT
TGCAATTGCATAACATTAAGTTCAGAAGATAATCTCCATAAGAATGGAATG

30   TTATACACTGTGTCTTTAGGATTGACCTTCCATAAAAGATACATACTTTTT
TGAGACGCACTTAAATCATATTGATACGATTCATTTGCTTTGGGAATCACT
TCATAGACATCATTTTGCAGTTCTTCAATTTGTTGTCCAAGTTGCTCTACA
GTAGGCGATTTCATTAAATCACCTACTTTAAGACGTTTTTTTAACCTTTCT
TCAATACGGTTTACAACTAATGTTGCTCTAAGAGAGTGTCCACCTAGTTCA

35   AAGAAATTATCTTTAACACCTACCTGATCAACATGTAAAATCTCTTCGAAA
ATACGGCAAACTGTGCGTTCAATATCGTTACGTGGTTCTACATAATTTCTA
TTATTCTTTAGATTAATTTCAGGTAATGCACGCACATCTAATTTCCCATTC
ATCGTGATAGGTATACGATCCACCTTCATAAAATGCACAGGTATCATATAT
TCAGGTAATGTTTCACTTAAAATATCTTTTAATTGACCTGTTGATTTTAAT

40   TGCGATGCTTCATAATATGCCACTATTTGTTTATCTTGGTCTTGCTCTCGA
ACGATGACTACAGCTTTATTAATATCACGTATAGCTTCTAATGCTTTTTCA
ATTTCTGATAATTCTATTCTAAAACCGCGTATTTTAACTTGCTTATCGATA
CGACTAATATAATCAATATAGCCATCTTCTTGAAGACGAACTAAATCACCG
CTTCGATAAAGCATTTCATTGTTAAAAGGTGACTGAATAAAACGTTCAGCA

45   GTAAGTTTAGGTTGATTTAAATAACCTTTTGCTAAACCTGCACCACCAATG
CACAATTCACCTGGAACACCTACGCCACAAATACGATTACCTTGCATGACA
TAAACTGTCGTTCCACTAATAGGTAAACCAATAGGTATACGTGAAGGCATC
TCTTGTGGAATCGCAAAAGTTGTAGTAAATGTTGTATTCTCTGTCGGTCCA
TAACCATTGATTATTTGAGGATGACACTCACGCGAATTTAATAAGTGAACC

50   CATTTAGCATTTAACACTTCCCCACCAATAAGTAAATAAGTTAAAGATTCT
AGTGCTTCGATACGTTCGCTAGCAATTTGATTAAATAAAGATGACGTTAAC
CACATCGTGTTGACTTTATTTTCAGTAATAGCTTGATCTAACAATTGAGGA
TTTAACAACGTATCTTTAGATGTAATGACTAACCGTCCACCATTCAATAAT
GGACCATATATTTCAAAGGTTGCTGCATCAAAAGCTACTGTTCCTGATAAT

55   AAGACGGTTGTATTTTCGTTCAATTCGACATAATTTGGATTGTGTACTAAG
CGATCAATTCCTCTATGTGGCACCAGTGTCCCTTTAGGTTTACCAGTCGTT
CCTGATGTATAGATGACATAAGCGATATCTTCTGAACAATTAATGCCTCTC
GGGTTATCAATATCATGTTCTCTTGAATCAACTATCAATTCTATATCCATT
TGAGGTAAACCTGATTGAAATGATGTACGATATGTTACAACCGCTTTAGGT

```
     TTTGCGTCCTCAATAATATAATTCATTCTTTCTTCAGGATAATCCGGATCA
     ATTGGTATGTAGCCTGCACCAGCTTTCAAGATCCCTAACATTCCAATAATC
     ATTTCTAAGCTGCGTTCTGCTATTAATGCCACCATATCATTAGGTTCAATC
     TGATGATTTAAACGCAATCTATATGCTAAAAGATTCGCACATTGATTTAAT
5    TGATGATATGTCATCGATCGTTGTTCATATTGCAACGCAATACTATCGGGA
     TGTTGGTGCACTTGACGTTCGAATCTTTCAACCACGGTTTTATAAGTGACC
     TCACTATAATTAATGTCATTGTTTTGAAGATTGATGTCATTATATTTAGCT
     AAATCTGAGCGATCACTGAGTTTTATATCTTTAATTAATAATGATGGATTT
     TGTGTAATTTGCATATAAATCCCATTAATTAATTTGACTAACGTATTTACA
10   GTTAATTCATCATAAACGTTGTCATTATAAACTATATCAAAACCACCCTGC
     ACATGAGGATAAAACTCTAAATCCGCGAATGATGATGCTTCATTGTATAGA
     CGATGTATATGATGACATATATGATTAAGCTCAAATTGTTCTTGTACTTTT
     TCAATATGAATCATCGTTTCTACATCTATTTGTACTGCTTTAGGTTGAACA
     ACAAAAGACGCACCACACATTTGAAGCTCTTCAAGTAGCGCGGAACACTCA
15   TCCACCATATCTTTTACGGCGTCACTTTGACTAATACTTAAATTAAGTGGG
     GCAATGTCTGTGTTTAAAACCATCATATTCTCAGTATTTTTATTATCTATC
     GAAAAATGGATGCCTAACGTCACATCAGAAGACTGACTTATAAAATGATGC
     GCTAAGTAAATACTTACAATCCACGCATCAATGTCTAAAGATGGTAGTTCT
     GCATGCAAAGAACACATTTTTTTCTTTTCAGATGTATTCTTTACAGGTATG
20   TAAGCGTAATTATCACCGTCTGCATTCAATAC


     Sequence 3431
     step.1003a07.cons.ok
25   AACATTAAGATACTCCCCTTGCTCATTTACACCCAAACTAGTGATAATTGG
     TATTATGTCGCGTGATAACATTTTTTCTAATAGCATCGTGTTGATTTCTTT
     AACTTTACCAACGTAACCATAGATATTTTTATTTAAAAATTCGGCGCTTAC
     AAGTTCTTTTATCTTTTCTTTAAATTGAAAAAATTCTAAATTAGATTCTTT
     AAGTCTTTCTAAAAGTTGATGGCCAACTATGTTAATTAAAGCATCATATAT
30   GATAGGTAAATCATTTTTAGCAGTTACTCTCATGCCATCAATTTTAATAGT
     TGAATGATTGTTTTTAGTTAATAAATTACTGATGACTTGGCCGCCACCGTG
     AACAATAATTATTTTTTTGTTTTCAAGGTGCCAAGCATTAATTTGTTGTAT
     AAATGCGTCGTTTAAATTTTCTATAGCTATACCACCGAGCTTAATTACGAT
     GATGTTTTTCATATGTTCCTCCTAAGTGGTGTATAGTGCGTTGATTTTAAC
35   GTAGTCATACGATAAATCACATCCCCATGCTTGACCTTCACAATTCCCTTG
     ATGAAGGTCAAGCTGAATTGATATTTCTTCAGCACTCATTATTTCTTGAAT
     TTCTTCTTTATCGTACTTTACTGGTGAGGATCTTATTAATACAGGTATCCT
     ACCTATAAAAATGTCTACCTGATTAATATCAAAATATGTTTTAGCATAACC
     TGCAGCAGCAATAATTCTACCCCAATTAGGATCTTCGCCAAAAATTGCGGT
40   TTTTACTAAACTTGAACCCACCACACTTTTAGCAATCATTCTTGCAGCACT
     AGATTCTTTTGCACCTTTAACCGTGACTTCTATTAATTTAGAAGCACCTTC
     GCCATCCCTTGCAATACTTTTTGCTAAATCGGTCATAATATATAGAAGCAT
     CTGCTTAAATTTATAGTAGTCTTCGCTGTCTTTTTTAATTTCGTTATTATT
     TGTACATCCATTTGACATCACAAGCACCATATCATTTGTTGAAGTGTCACC
45   ATCTACAGTGATTTGATTGAATGTAACTTCAACCACATCTTTTAAAGCCTG
     TTGTAATGTTTGTGATGAGATGTTAGCGTCACAGGTTATAAATGCTAGCAT
     TGTAGCCAAATTAGGATGTATCATTCCTGACCCTTTTGCTACACCTGCCAT
     CGTTACTGTATCGCTACCAAATTCTTCGTTTACAACGCATGTTTTTGTATG
     AGTATCCGTTGTTAATATCGCTTTTGCAAAGTCATCAGCATTACCGTTTTT
50   AACTAGTTTGGAAAAGCCATTCTTTAGAATAGACATTGGCATCACCTTTCC
     AATAACACCAGTAGATGCGACACCAACATATTCTGGTTGAATTTGTAATTT
     ATTTGCGGCCAGTTGTTGCATTTTAAAAGCATCTTTTTCTCCTTGTTTACC
     AGTACAAGAATTAGCAATACCTGAATTAACAACAATAGCTTGCATTTTACC
     ACTTTTTTCGATGCTGTTTTTTGTTAATTTTAATGGTGCAGCAATGACCTT
55   ATTAGTTGTAAATACACCAGCTACATTTGCAGGTACTTCTGAAACAATCCA
     ACCAAAGTCTAATTTTTTTCTTTTTAAAGCCAGCGTGCAGACCATCAGCTGA
     AAATCCAAGAGGACTTGCAATATTTCCCTTAATTATATTCATAATTTAGTC
     CCCCTATATTAGATATATAAAGGCGCCAAAATTAAACCATCAGTTTCGTCG
     AAATTAAACATCAGATTCATGTTTTGAATTGCTTGTCCAGCGGCACCTTTA
```

```
      ATTAAATTATCAATAACTGATGAAATTGTTAAAACACCCGTGGTTTCGTTA
      TAAACAAAACCAATATCAGTGTAATTTGTTCCTATTACTTCATTTAACTGT
      GGTAATCCATCTTTAATTCTTATGAAAGGCTTATTTTTATATACATCCTTG
      TAGGTAGATTCTATTTGATTTATTTTAACTCCGTTTTCAAGCCGTGTGTAT
5     ATAGTTGCAACAATTAATTAGAATCTAGTGATTGATTGTCAGATTGTACCT
      TAGCAAGTGCTTTTTCAATCTTATCACTACCTTGTAGTGAACGTGTAATTG
      CTTGTATATCTTGTTTTTCATCTGTGACATGTTGATCTTTTGAATCCCTAA
      TTTGATCAGTATCTGATGCTATACGTTCTAATGTTGAGCGACGACCATCAT
      TACTTTTATCGTCTTTACTATCACTATCGGATTGATTGTTGTCGTACTAG
10    ATTCTTCTTTGTTATTTTCGTTTTTATTACGATGGTGATGCTTGCTGATCG
      CTTTGTTCAGAAAGCTTATTATTAGAGTCAGACTTAAAATCTTGTTGGATT
      TTATCTAATTTATCAATAACCGACTCTTCAGAATCATCAGTAGTTGCATTG
      CGTGTACTAGCATTTGTATCGTGTACTTTATTATTTGAATCATCTTTTTCA
      TCGGTTGTTGAGTCACTAGGTTTAGAAGCACTATTTTCACCTTGTTGACGG
15    TTATCAACATTTTTAGCAATTGAATCTAAATCCGACAGTATTTTATCTAAG
      TCTTTAGTCATATCATCATTATCTTTTTCAGATGTATCGTCATCTTCAGTT
      TTTGATGATGTGGCACTAGAGTCTGTCGATTTTAAATGATTATTTTTATCT
      TTTGCATCATTTATAATACGGTCAATACTTGATTGATTATCGACTTTATCT
      AAACGTTCTTCAAAACTTTTGATATCTGCGTCAGAAAGTTCTTTCAAATCT
20    TTAATATCATCTTTAGCTTGTTGGATTTTGTCAGCGATGGAGTTACTATTA
      GAAGTATTTTCAACATTTGATGATGTACTTTCTGAATCTGCAAAAGCATCT
      TTTGTGAAGCCATAACCAACCAGCGAACCACTAAGTAACAACGCAGAACAT
      GACAGCTTGAGGTAGACACCAATTTTATCTGTTCTCTTCATTTACACATAC
      TCCTTTAATAAAGTTACCTGTATTATTCTTTAAATAGATTTATATGATTTA
25    ATCAAAAATATACATATCACATTATAATACATAATTTCATAGATTTGTAGA
      TAATGAACCTAAAATAACAAATTTAAAAAACAAATTTCTTGTAATTATTAG
      TAATTTGGTATTACACTTCTTT


30    Sequence 3432
      step.1003a08.cons.ok
      TGGATGAACTTGTTGTCCTTTAATAATATGACTTGCTTCAATAAGATCAGA
      TAATCTTGCATTCGTGCATGAACCTAAAAAAACATAACCTAATTTTATATC
      TTCGGCTTTTTGACCTGGGTGAAGTCCCATATAGTCATAAGCACGTTGGTC
35    ATTTGCATTTTTAATTTCTGGGAATGGATTACTAAAACTAACTCCCATTTC
      TGGGTTAGTTCCCCAAGTTACTTGAGGTTCTAAATTTGTTACATCAAGTTC
      AATAACTTTATCAAAATAGGCATCATCATCAGAATAAAGTTCTTTCCACCA
      CGCCATAGAACTATCAAAATCAGTAGCATAAGGACGACCTTTTACGTAATT
      AAAGGTTGTTTCATCAGGTTGCATTAAACCATACTTTGCTCCTGCTTCAAT
40    AGCCATATTACAAATCGTCATACGTGCTTCCATGGATAAATTTTTTATAGT
      TTCTCCTGTAAATTCAAGAGCATAACCAGTACCGAAATCTACACCATATTG
      ATTGATTAAATATAAAATAATGTCTTTAGCATAAACTCCAGCTGGTAAAGA
      ACCATTTATATTAATTTTCAAATTTTTAGGTTTTGTTTGCCATAATGTTTG
      TGTGGCAAATACATGTTCAACCTCGCTTGTTCCTATCCCAAATGCGATAGC
45    ACCAAATGCGCCATGTGTAGCAGTATGTGAATCTCCACATACAATAGTTTT
      TCCTGGTTGAGTTAATCCAGTTTCTGGTCCAACCATATGCACAATACCTTG
      TTCGTCAGAACCCATATCAAAAATATGTACACCAAAGTCCTTAGCATTTTG
      TTGTAAAGTTGTAATTTGTTTATTAGCAATTTCATCTTTTATATTAAAAAT
      ATCAATTGTGGGAACGTTATGATCTAAAGTTGCAAAGGTTAGATCAGGTCT
50    TCTGAGTTTACGATTTTGTATTCTAAGTCCTTCAAACGCTTGAGGAGAAGT
      GACTTCATGAATGAGATGTAAATCAATGTATAATAATTGTGGTTCACCTTC
      TTTTCCATGAAGCACATGTTTTTTCCATACTTTATCAAACAGTGTTTGACC
      CATAATTTTCTCCTCCTCTTATAGATATTTTTCTTTAAGCAATTTAAAAAT
      TTCTGAAGTTCGATATTGTCCACCTAAATCTGCAGTTGTCTTATTAGATTG
55    AATAAACGAGTAAACAATTGACTCAAGTTCGTTAGCAGCATCATTTTGATT
      TAAACTTTCTCTTAAGCAAAGTGCTAAAGATAGAACCATACCAAATGGATT
      CGCTTTATCTTCATTAGCTATATATAAAGAACAATGTGAACTGACCTGTTA
      AAAGTCGTGCAACAAGTCCTCCACCTAAGTTGTGGATAGCTTTTACATTTC
      CATCATTGGTAGCATTGATAGTTGTACCATATGGAGCGCCATAGTCAATAC
```

CAAAGTGAGCACCACCATTAAAACTATAACCAGGAGCACCACCATTAGGAT
AATACCCCGTAGTAATTGGGAATTTAGTGAACGAAGAACCATCTCCACCGC
CTGCAGAGTCTAAACCTTCATTAAAGAAGTCTTTAATACCATTTTTAAGTT
TCTTAATCATAGCTGTCATTAAATCAAATGGTAATGGTGCATTTTTCATAA
5   AGTCTAGGTTAAAACCAACTTTCTCAAAGACTTTATTTACCAATTTCATAG
GATGACCTACATAATCAAAAACATCTCCAATACCCTTGATAACAGTGCTAG
ACACATCTTTAGCTTTATCAAGTGCTTTTGTACCTATGTCTTTTGCTTTGT
CAAACGTAGTACCACCTATGTCTCTTGCTTTACCACCAATGTCTTTTACTT
TATTAACATCATTTTTTACACCAGAGCTAATAACATCTAATATACCGTCTT
10   TCTTTTTAGTACCTTTACTAAATTTAGGTAGTTTTTTCTTCTTAGTGTCAT
ATCCTGAATTAGATAAAATAGCATGTGTTTGCGCGCCACTATATACACTTG
AGCCTTTAGGTAAGAATGTTGTTGTATCTTTGTTAGGTGTTAGTGCCATTC
TTCCGTTAGGATAACGGATCATTTCATTTCTAAAGCCACCAGGTCCATTAC
CTCGACCTCTATCTCCTACAGTAGCCATTGTACCTTTATTGATTTTACCGT
15   TGGTAATATAACTTTGTGTATGCGTAGATTGAGTACCTGTAGATAATTTAA
TCTTAGGTATTTTATCCATACCTAATTTATCTGCGACCCAGTTAACACCTT
TGATTAATCCATTAAGGCCTTTTTTAACGCCTTTAACCATTCCACCAAAAA
GATTCCCAATTTTACCGGTAACTGTTTTGATACCGTCTCTCATTTTATTCA
TGGTGCCCATAACTTTCGATTTCATGCCATTTACTATAGATACAGTAGTAC
20   TTTTAATTCCATTCCATTTCTTGCTCATGAAGCCACCTACAGCATTCATGG
TGTTATGAGTACCTTTTTTAAGTGATCCCCAAGCACCTTTGACGCCTGACC
ATAGAGCTTTCGCTTTATTAACAGTACCTTTTTTAATACTGTTCCATTTAG
AACTCATGAAACTGCCAACTGCTTTAAATATGCCAATTGTACCTTTTTTGA
GTGCATTCCATGTATTTTTAACTCCAGACCATAATGCTTTTGCTTTATTCA
25   CTACTGATTTTTTTATAGCCGTCCAGATTTTAACTGCGGCATTCTTTACAG
CATTAAATATTACAACAATACCTTTTTTTTAATGCGTTAAATACAGATAGAA
CACCTTTGCGCAAAGCTCGAACGATTCCTAACACACCATTTTTTAATGCAG
TCCAAACTTTAATAGAGAAACTCTTAATAGCATTAAATATCGTAACTACTA
TGCGCTTAATAAGGTTGATATTAAATTTTACTTGCGCAACATATGCTTTAA
30   TAAACTAAGTAAATTAGAAGTACAAGATCGAACACAAGCAGTAATATATGC
GTTCCAGCATAATTTAATTCAATAAATTATTAAAGGCGAAAGTAAAGATAC
ATCTATCGATACTTTCGCCTTTCAATATTATCAGTCGCTTGTGTGATGTTC
TTC

35

Sequence 3433
step.1003a11.cons.ok
TTTAATTGTTCTACTTCTGACAGCTCTTCTAATCCTTTACCGTAGGTATAT
TGTTTACCAACTTGTTGTGAAAATCTATAACTTTCCCCATTTCGATACCAT
40   CGCCACCAAGTTTTCACTTGCGTCCTATTTCTAATATTTAATTCTTTCATA
ATTTCTTTTGTTGAAAATCCTGCTGCTTTCATTTCAACTGCTTTATACTTT
GTTTCTACTGAATAAGAAGCTCTTTTTCATAGAAAAAAACCTCCGTATGATT
CATTTTAATATGAATTCAACGAAAGTGTTTTTTATATAATTCCCACAAATTG
GGGTCAGTCTAATAATGAATTTCGTGATTGTAGGCTTTTTATTTTATATAT
45   TCATTAATTTCTCTTCTTGACGGATACGTACAATTAAAAAAATAAGCATATG
GAACTAATAAGAGTGTTGTGTATGTAGCATTTGTTAGTAGTAATACACCAA
TTAATTCGGGAATGATATTTAAAAAAATAGTTTGGATGTTTCGTTACTTTAT
ATAACCCTGACTTAATAATAGGGTGATTGGGCAGTATAAATAATTTCAATG
TCCAAATACGACCTAATGTCTTAATAACTATAAATAGCATGATATAAGCGA
50   TGATCAATATGATTAAGCCAATGCCGTTAAGTAGACTAAATGTATCCTTAC
GAATGAATGCTTCTATAGCTGCACTCATGTAAATTAATACATGCGTAATGG
CTAGATATTTTGAATTTTTCACACCATATTCCACCGCGCCCTCTACCTTTA
GCTGTTTTGCGTGTTGCATAGATATCTTTAAGCTGATGAGTCGAATACAGA
AAAAGATAAATAAAATAGTTAAAATCATGATGTCCTCCATTGTTTGTTATA
55   TATATATTCTTTTGCTCGTACAATACAAAGTATTTTATATAGAAAATGTTT
CAAATGCCATTAACTATTTATTAACTTTAGCACCAACTCATAACTCTTCTA
ACACCTCACATATTAAGTTCAGTTTTTCGGATAATTTAATACTTTTAAGGA
TATTAAGCGCTTACATTGATGTGATATATTTTTTTTAACGAAGATGATAGG
GGTGTTAGGGATGAACTTTTTTGATATTCATAAAAATGCCAAACAAAGGGAT

```
      ACCATTAGCTGTACAACGCAAATTATGGCTCAGAAACTTTATGCAAGCGTT
      TTTTGTCGTATTCTTTGTTTACATGGCGATGTATTTAATTCGAAACAATTT
      TAAAGCGGCACAACCGTTATTAAAAGAAGAAATCGGATTAACAACATTAGA
      ACTAGGTTATATAGGATTAGCGTTTAGTATTACTTACGGTTTAGGAAAAAC
 5    AATACTCGGTTATTTCGTTGATGGGCGTAATACGAAACGTATTATTTCCTT
      CTTATTAATATTATCTGCGATTACAGTACTTATTATGGGATTTGTATTAAG
      TTATTTCGGTTCTGTGATGGGGCTATTAATTGTATTGTGGGGGCTTAACGG
      TATATTTCAATCTGTGGGTGGGCCTGCAAGTTACTCAACGATTTCAAGGTG
      GGCGCCTCGAACAAAGCGCGGTCGTTATTTAGGCTTTTGGAATACATCACA
10    TAACATTGGTGGTGCTATTGCTGGTGGTGTCGCACTTTGGGGCGCGAATAC
      ATTTTTCCACGGTAATGTGGTTGGAATGTTTATTTTTCCTTCCGTCATCGC
      TTTAATCATTGGGATTGTGACATTATTTATTGGTAAAGATGATCCAGAGGA
      ATTAGGTTGGAATCGTGCCGAAGAAATTTGGGAAGAGCCTATCGACCAAGA
      AAACATTGATTCTCAAGGTATGACTAAATGGGATATCTTTAAAAAATATAT
15    CCTTGGAAATCCTGTGATTTGGATTTTGTGTATCTCTAATGTTTTTGTATA
      TATCGTGCGTATTGGTATTGATAACTGGGCACCGCTATACGTATCAGAGCA
      TTTACATTTTAATAAAGGTGATGCGGTGAATACTATTTTTTACTTTGAAAT
      AGGTGCATTAGTAGCTAGTTTATTGTGGGGCTATATCTCAGATTTATTAAA
      AGGTCGTCGTGCGATTGTAGCGATTGGATGTATGTTTATGATCACCTTTGT
20    TGTACTCTTTTATACCAATGCAACAAGCGTGACAATGGTCAATATTTCTCT
      ATTTGCATTAGGCGCTTTAATCTTCGGTCCACAGTTACTCATTGGTGTATC
      TCTGACTGGCTTTGTTCCTAAAAATGCAATTAGTGTCGCTAACGGTATGAC
      AGGTTCATTTGCATATCTATTCGGGGATTCAATGGCTAAAGTGGGTCTGGC
      TGCAATCGCTGATCCAACACGTAATGGTTTAAATATTTTTGGGTATACGTT
25    GAGTGGTTGGACAGATGTCTTTATTGTATTCTATGTAGCTTTATTCTTAGG
      AATGATATTATTAGCCATTGTTGCTTATTACGAAGAAAAGAAAATTAGAAA
      ATTAAAAATTTAAGATTAAGTGAATTTAAATATATACTCCCTACTATAAAT
      TTTATCATTATGGTAGGGAGCTTTTTCGTGATTGATTTGCACTATTTTGAT
      TGCTTATTATAATCACTTGGTGACATATGTAAATATTTTTTAAAATGATAG
30    CAAAACATTTTATACTCAGAAAAACCTACTTTTTCAGCAATTTCATAATGC
      TTGTAGTGCTGGTCTAAAAGATGTAATGATTTTAAAATACGATAACGATTT
      AGATAATCAACTATCGTAATGCCTACATGTTCTTTAAACGTCCTCATGGCG
      TATGACTCACTTACATCAATAGGATTAATTAAGTCAAGAACAGTCACTTTC
      TTGTGATAATGTTGCTTGATTTGAGACAAAATTTGATTGACATAATAGTCA
35    TCGTAATCAATTTTTAATAATGGTTGAAAGGCAGTATGATATGCCGCGTCA
      TCATTGGTAGAATGTGGGCGTTCTAATAACCTTTGAACTAATATGTCTAGA
      ATATGCTCTAACTGAGTGTGGTCTACTGGCTTTAGTAAATAATCAAGAACA
      TGATGTTGTATACCGGCTTTCATATATTCGAAGTCGTCATAACTCGATAAA
      ATGATAATCTGGCAATCAAGGTCTTTGATGTCATCTAGTAGGTCAACACCA
40    TTTTTTCGAGGCATTCGTATATCAGTAATGACGAGTTCAGGTTGATGTTGG
      CGAATTAAAGACAATGCTTCCACGCCGTCTTTGGCAGTATAAACAGTGGTG
      AAATGATAGTCCTCCCATGGAACCATTTGCTTTAAGCCTTCTCTTATAATC
      CTTTCATCATCACAAATAACTACTTTAAACATCTTTATTCCTCCTAGACAA
      GTGGTATTTGGTAACACATTAATGTCCCTTGCTGGCTTCTTGAGAAAATGT
45    GCAGACGTGCATATGTTCCATATTGAATCATGGCTCTATTATGTAAATGAT
      TTAAACCTAGGTGTGTCGTATCAAAAACATCGTGATGAAGGGATTGGCGCA
      CGTGTTCTAAATGTGATGGAGACATACCGATGCCATTATCATGAACCAGAA
      TATGTAATTTGCGCTTCGTAAGTCTGATACGAATTGTTATCTTTAAAGGTT
      CACTATCACGACCATGCTTGATGGCATTTTCTACGAGTGGTTGAAGCATCA
50    TCTTACCAATTGTTTGATGTTGTACACCCTCAGTAGCATCGATGTAAAGCT
      GTATCATATCATCGAAGCGGATATTTTGTATGGCAACATACTGCTCAATGT
      AGCTTAATTCTTCTGCTAATTTGACTGTGTGCGATGCTGTGCGTAGAGAAT
      AACGTAGCATTTGTGATAGTTGTTGTATGACTGTTTGTGCTACTTTGGGTG
      AAAGAGGAAT
55

      Sequence 3434
      step.1003b04.cons.ok
      GCGCGGCAATTGAACAAAAGAAAAAAGGTATTGAAACATTAGTAATTGAAA
```

AAGGTAATGTTGTTGAATCAATCTATAATTATCCAACACATCAGACTTTTT
TCTCATCAAGTGATAAATTAAGTATCGGCGATATTCCTTTTATTGTTGAAG
ATAGTAAGCCAAGACGTAATCAAGCGCTTGTATATTATAGGGAAGTCGTTA
AACATCATCAACTTAACATACATCCATTCGAAGAAGTTTTAACAGTTAAAA
5 AAATAAACAATAAATTTGCAATTACAACTACAAAAGGTGTATATGAATGTA
AATATTTAACTGTTGCTACGGGTTATTATGGTCAACATAACACTTTAGAAG
CGGAAGGGGCAGAATTACCAAAAGTATTCCATTACTTTAAAGAAGCACATC
CGTATTTTAATCAAATGTTGTTATTATTGGAGGCAAAAACTCTGCTGTTG
ATGCTGCCTTAGAATTAGAAAAAGCTGGTGCTAATGTAACTGTTTTATATC
10 GTGGCGAACAGTACCCTAAAGCAATTAAACCATGGATATTACCCAATTTCG
AATCATTAGTCAATCACGAAAAAATTACGATGGAATTTAATGCGACAGTAA
CCAAAATTACCGATCATTCAGTGACTTATGAAAAAGATGGTCAACTTATAG
AAATTGATAATGACTACGTTTTTGCTATGATTGGTTATCATCCAGATTACG
ATTTCTTAAAAACAATAGGTATTGATATCCATACCAATGAATATGGAACTG
15 CTCCTGTTTATAATCGAGAAACATTCGAAACAAACGTCGAAAATTGTTATA
TAGCTGGTGTTATTGCTGCGGGTAATGATGCAAATACTATTTTTATCGAAA
ATGGTAAATATCATGGTGGTGTCATTACACAAAGCATTTTGACAAAAAAAC
AAACACCTCTTGAAACATAGAGCTTACAAAAAGCTTGGGATACTAAACAAT
GTCCCAAGCTTTCTTAATGTATAGTCAATCAATGACTCAATTAAAATGCAC
20 ATATTAATAATAGGTAATCCTTAATTATAGTATAATTGCACTAAAGAAGTT
TCATAGACAATGCATACAACTAATAGAAATCGGAGTGTGACAATGAACTTG
ATACGAATTTTGTCCCACTCTTTTTAATTACTCTTCGAAATATCGCTCTAA
TTGATTTTGAGTCATGTCTTGACTCAAACCAACTAGTAATTTTAGCCTTGC
CTTTGGTCCATTTAAACCGTTCGAAAAAATAACACCATTATTTTTCAAATC
25 TGCGCCACCACCTTCATAAGCATATACAGGACTAACAATACCATTGAATGA
TCTAGATACGAGAACTAGAGGAATGTTTTTCTTTAGACATTGCTGTAGTCC
ATTAAGACAACTTTTTGGAAGGTTACCTTGTCCTAGTGCTTCGATGACTAT
ACCATCAACGTGTTGTTGTGAATAAAATGATAGTACATCATCTTCCATACC
CATGTATGCTTTTACAAGTGGAACACGTAAATTTACATCGATATATTGGTA
30 GGTAGTTTGTCTGTAAGGATGATGATAAAATTGTACTCGATTCTTGGTAAG
TACACCTAGAGGCCCCTGATTAGGACTTTGAAATGTATTAATATTCGAAGT
ATGTGTCTTTGTCACATTACGAGCAGTGTGAATCTCATCATTAAATACGAC
CATAACACCTTTATGATTAGCCTCAGATGAAGAAGCAACCCTTATAGCAGA
AATAAAATTATAGAGACCATCGGAACCAATTTCATTGGATGATCTCATTGC
35 TCCAGTAATTACTATAGGCTCTTGAATATCAATTAATAAATCTATTAAAAA
AGCTGTTTCCTCAAGTGTATCTGTTCCATGAGTAATGACAAATCCATCATA
TATATTTTCTTTAGAATATGTAATGATTTCGTCTCTTAATCGCACAACATT
CGAAATTGTCATATGCGGCGAGGGTATATTTAAAAGATTGATTTCGTCAAC
CTCTGCATATTGACTAATGATATTTTGATGTTGTGATATTGGATTTTCTTC
40 ATTCGTTATCACTTTATTAGTTTGATCTTGTGACATACTTATTGTGCCACC
AGTATGTATGATAAGTAGACGTTTCATTACGAATTCCTCCTATTTATTCTT
TACTTTTTTATGATAAAATAATATCTATAAAACAACAAGGAAGGTTATCAA
ATAATGAGTTCAATTAATATTGCACTAGATGGCCCAGCTGCTGCAGGTAAG
AGTACAATTGCTAAACGTGTAGCCAGTCGTCTATCAATGATATATGTTGAT
45 ACAGGAGCAATGTATCGTGCCATTACATATAAATATTTACAAAATGGCAAA
CCCGAAAATTTTGATTATCTGATTAATAACACTAAACTTGAGCTTACTTAT
GATGAAGTAAAAGGGCAAAGAATCTTACTAGATAATCAAGACGTCACTGAT
TATTTAAGAGAAAATGATGTAACACATCACGTATCTTATGTTGCATCTAAA
GAACCAGTGCGTTCATTTGCAGTGAAAATACAAAAAGAATTAGCTGCTAAA
50 AAAGGTATCGTTATGGATGGCCGAGATATTGGTACAGTTGTATTACCAGAT
GCCGAATTAAAAGTTTATATGATTGCATCTGTTGCTGAACGTGCTGAACGT
CGACAAAAAGAGAATGAGCAACGTGGCATTGAATCAAATTTAGAACAATTA
AAGGAGGAAATTGAAGCACGAGATCATTATGATATGAATCGTGAAATTTCG
CCATTACAAAAAGCCGAAGATGCTATTACACTTGATACAACTGGCAAATCT
55 ATAGAAGAGGTAACAAATGAAATATTATCTCTACTTTAAATGTTAAAATTA
AACTTTTTATATTGTTTATTGTAAACATTTAAGCGGAGTAGAATTAATAAT
CAATGACTCGAAAATATTGAATATATTAGTATTGATTTATAGAAAAGCCAC
GGACGCTCTGTCTCTGGCTTTTATTTTTGACCATTGAGAACTATGTAAATA
AACTTATTTTTAACCAAGTCCCCTCTATAGAGAAGTAATACTATGTTTATT

ACCAATTTTAATATATCTTTTTCATTTTTATTCCTCATATTATTATGAGCT
TAGGGTTAATTGCCATAAAATTTATGTTACAATAAAAATAACAATGTTATT
AAATATACTACAAACTAAATTGACCATCACTTAACTAGGTGCTCTGTTAAA
CTTTTGTTATAAATTAAAGATAGAAGCAACCTATTCATATAAAGTGGTATC
5 TCTTGACAATTTTACTCGTTTATAAGATGTTATAATTATGTAGTGTATAAG
GAGGCATACAAGATGACTGAAGAATTCAATGAATCAATGATTAATGATATT
AAAGAAGGTGACAAAGTCACTGTTGAAGTTCAACAAGTAGAGGATAAACAA
GTTGTTGTGCATATTAATGGTGGCAAATTTAATGGAATTATTCCTATTAGC
CAGCTTTCAACACATCATATCGAAAACCCTAGTGAAGTTGTAAAAGTCGGT
10 GATGAAGTCGAAGCATATGTCACTAAAATCGAGTTCGACGAAGAAAATGAT
ACTGGGGCATACATTTTATCAAAAAGACAACTTGAAACTGAAAAATCTTAT
GAATATTTACAAGAAAAACTAGATAACGATGAAGTGATTGAAGCTGAAGTT
ACTGAAGTAGTTAAAGGTGGTTTAGTCGTTGACGTTGGTCAAAGAGGGTTT
GTACCTGCTTCTCTAATTTCAACTGATTTCATTGAAGATTTTTCTGTATTC
15 GATGGTCAAACAATCCGTATTAAAGTGGAAGAACTTGATCCTGAAAACAAT
AGAGTCATTTTAAGCCGTAAAGCTGTGGAACAGTTAGAAAACGACGCTAAA
AAAGCTTCAATATTAGATTCTTTAAATGAAGGCGATGTTATTGATGGTAAA
GTTGCTCGATTAACTAACTTTGGTGCTTTCATTGATATTGGTGGCGTAGAT
GGTTTAGTTCACGTTTCTGAATTATCTCATGAACATGTTCAA
20

Sequence 3435
step.1003b05.cons.ok
ATCTACTATTTTCTCAACATCTTCTTTTGTATGAATGCCACTCTCTGAAAT
25 GTAGCAGCAATTAGACTTTTTAAACTTAAGTAATTTATTTGTATGTAGAAC
ATCGGTTTCAAATCGTTTTAAATCACGATTATTAACACCAATAATTTTAGG
GTTAATTTGGTGTGCACGTTCAAGTTCTCTAATTGTATGAACTTCTACTAG
AGCTTCTAAATTATGGTTTGTTGCATATGAATACAATTCTTTTAATTGGTC
ATCACTTAAAATATTTACTATTAATAAAATAATAGATGCACCAGCTCGTTT
30 TGCAACATCTATTTGAATTTTATCAATAATAAAATCTTTACATAAAACTGG
TAACGATGTTATCTTTGATAACTGATTTAATCGTTCAAAACTACCGCCAAA
GTATTTTTCATCAGTTAATATTGAAATAGCATTAGCACCATATTTTTGATA
ATCTTTAACTTGTTGAACAAGATCACGTTGCGGTAATTGAGGTACAGATGG
GCTTTTCGATTTTATTTCAGCAATAACTGATAATGTTCTATCATAGTTAAG
35 TGAATCAATCAGCTTTCTCCTCTTAACATTTCCGTTATCTTGTAAAATTTC
AAGTTTTTTATCATAGTATTTACGCTCAAGCAAAGTTTTTTTATACTCAAT
AATTTCATTTAAAAATAGTCATTACTTAACCTCCCATTTTTAAATATTGTTT
CATTGCCATACCTGTATCAATGAGATATTTCGCTCTCTCTACGCCATGTTT
GATACTTTCCACTTGCTCAGCAACATATAAAGCAATTCCAGCATTTAACAA
40 AACTACATCTCGTTTACTTGAGTGATCCGTGCCACTTAGGATATTCAATGC
AATTTGTTTATTTGTTTGAGGTGAACCACCTATCAACGTGTCATTATTTGC
ATAAGCTAAACCGACTTCTTCTGCTTTTAAACTATATTTTTTTAATGCTCT
TTCGCTGCTAACTTCATAAATGATATTTTCACCAGAAAGCGTGGCCTCATC
CATCCCATTTGCACCATGAATTAAAATTGCTCGTTTTCTACCTAAATCCTT
45 TAATGTTTGTGCTATATTTTCAAGTTGTGAAGCTTCATATACCCCCATCAC
TTGATAAGTTAATTTGAAAGGATTAATTAATGGTCCAATCAAGTTAAAAAT
TGTAGGTGTTGCAATCGATTTTCTAATTGATTGAAGCTTTTTCATCATTGG
ATAAGAATCAGTTGCACTTATGAATGCTAATCCTTTCAAATTTAATTGTTG
CTCTACTTCGTTCATTTTGTTTGTTTTTATATTCATTTCATGTAATACATC
50 TGTACTTCCTGAATGTGAAGTAATACTTTTATTACCGTGTTTAATGACTGG
CACTCCTGCACTTGCTACAACAAAAGCTACAGTTGTAGAAATATTAAAGCT
ATTTGATTGATCTCCACCTGTGCCACAAACACACATAGCTTTATTATAAAA
TGGTTGGTTTGGATAGTTTGTCTGGATAAAATATTCAACTAAATACGTTAG
CTCATATTGACCCATGTCTTTATTTGTATAAGCTTTCAATAATTCAACCTT
55 TACATTGGTTTCTATATTTGAATCAAACAGTGTAACAATAAATGATTGCAT
ATCTTTTTTAGATAAAGATTTATTTTGTTTAATTTTCTCAAGAAGGGTCAT
TTTCAACTTCCTCTCGTTGCTATATTTATAAAATTAATAATGATTTGTTTC
CCATTTTCTGTAGCGAATGACTCTGGATGAAACTGTATGCCATAATGTTTT
CTGTTTTTGTGTTGGAAAGATTGTATACTATCATATGTTTGGCCAGTGATT

AATAAACTTTCTGGAAATGAGCATGGATCACTAATTAATGAATGATAACGC
ATAACTTTGAAACTTGACTTAATTCCTTTGTATAGTGGTGACTGATTAACA
ATAGTTAAAGTATCTACTTTTCCATGTTTGACACAACTACCTTGTATAACA
TCGCCACCGTAATAACAAGTGAGCGCTTGTGCACCTAAGCAAACACCTAAA
5 ATAGGTAAATTTTCGAAATTTTCAATTAATTTAATTAATAATTGATTGTCT
AATGGATGTCCAGGTCCTGGTGAAATAATAACTGCTGTAATACCTTTCAGA
TTACATATATCTGTATCATCGGGATATTTAACAATCGTTTCACATTGTTTT
TCTACAATATCTACAAGATTGTAGGTAAAAGAATCATAGTTATCTATTATT
AATATCATGGAGTTACCTCCAATAAACTTTTAGCTTTAAGTTTTGTTTCTT
10 CAAGTTCTTTCTCTGGAATAGAATCATATACTACTCCACATCCTGCCTCGA
CACTGACTTTTTCCTCATCGATAATCATGGTACGTATAGCCAATGCAAAAT
CTAAATGATGATTACAGTTGATATACCCAACACCACCGCTATAGATACCTC
TTTTATAAGGATAAGATTCGTATATTCTCTGTATAGCTCTAAGTTTAGGTG
CACCTGAGACAGTACCCGTTGGTAGCAAACTTGCGATGACGCTCATAGGAG
15 ATAGATGGGGTTTTAATTCTCCAATAACTTCACTAACGATATGCATGACAT
GTTCATAACGTTCTATTGTCATTAGTTTGGTAATTTGTGAAGTGCCTGTTT
TACTTATTCGATGAATATCATTTCTTCCTAAATCTACGAGCATACGATGTT
CACTCAATTCCTTTTCATCTTTCATTAATGTCTTTTCATTATTTTCATCTT
CTTTTTTATTTTGACCTCTTTTAATTGTTCCAGCTATAGGATTCGTATAAA
20 CTTTTCCATCTTTTACCTTTACAAAACTTTCAGGAGAACTTCCTATTACAA
TCGGTACATCTTTATTAATATAATACATATATGGACTAGGATTTTGTCGCT
TTAAATTTTGATATAACTGAAAAGTTAATTGATGTAAATTGTGTTGAAAAT
GGTGTTTATAACTATAAATTCTTGAAGGAACTACTTGAAACATATCTCCTT
CAGTAATTTTCTTTTTTAAAATTCTAATAGTTTGAACAAATTGTTGCTCTG
25 ATATATTGGTGGTTATATGTCGAGGGAT

Sequence 3436
step.1003b06.cons.ok
30 ACAATTATATTTGATATAAGTATGGTTGTAATTGAAAATAAAAAATAGGAC
AGAATTCATGTTGAATCGACTGTCCCACTCTTGTGAATGATATTACTTTTA
GTGATTCTTTGACATAGCTATCTTATTAGACAATATTAAGTTTATTGACAA
TTATCATCAAAAGCATCTTTTAAATCCATAGCTATATCATTGATGTCGCGC
CCTTCAATATGTTCGCGAGGTATAAAATGAACTAAATGTTGACCTTTAAAT
35 AATGCGTATGAGGGACTTGAAGGAACTTGTTGGATGTAATCTCTCATTGTT
TGTGTTGCTTCTTTATCTTGACCAGCAAATACAGTTACTTTATGATCTGGT
TTCACTTCATTTTGCTCTGCAACTGCAACAGCTGCTGGACGTGCTAATCCT
GCAGCGCAACCGCATGTTGAGTTGATTACAACAAATGTTGTATCATCATTA
TCTATATTTTGCATATACTGATTGACGTCATCACTCGTTTCAAGACTTGTG
40 AATCCATTGTCTGTTAATTCAGCTCTCATTTGTTGTGCAAGTTCTTTCATA
TAAGCTTCGTATCCATTCATCATAAAATGCTCCTCTACTAATAATTGTTTC
TAATTGTACTATATAGAAAATTAATTGCTTTTAGCGAACGTGTTTTCTATT
TGTGGTAATAATTCATTCCAGTTAGCGTTATCTTCTTTATCAATTGATATA
AAATCCAAAACATAGAATATTGATTTAACACCTTCTATTTCAAATAATCGA
45 TTAATAAATTCAGGTTGACCTTCTTGAGCGGCAGTATATGTTGTAGATGAG
TTATCTTGACGTGGTTCACTCAACGAAACTTTCATTGTATTATGATTAGGT
GTTTCTGAAATTCCTATAATCTCCATGTGACATCCTCCTATTGTAACTCAT
TACCAATTTGTTTAATTTGTTTACCGACCGTACATTCATTTATACAAAAAT
GATGTGCTTGTGTTTTACCGTGAGTTTTTCTATTGTGCTTTTTTTAATGGAC
50 ATTGGTGACAATAAGTATTCATCAATTGGTCGATGTGTTGAATTGCTTGTT
GTTCTGTATTAGTGAGTATAGCTTTCACTTCCCTTAGCTATAGACTATTTA
TATTTATACCCATTTCACATGTTTGTATTATAGCATTTATAAGTTGGGAAT
TTGAAATAATCAATTTTAGTTTTGAGAAAACTGATAACTTAGACATTTGTT
TAAGTTATTATTTATTTACTCTTCTTTTTTAGCATTGCATTAAAGATGAAT
55 TTTAAGTAAAATAGTATAGTTAATTTTGGCGGTTTCTAAACACCCGCTTTA
ACAAAATTTAGGAGGAAAGTTATATGTATAATGAAATATTTGGTATTGCGT
CATTTATTGTTACATTCGCTTTAATGGTACTGATGTATCGCTGTTTTGGTA
AACAAGGACTAATTGCTTGGGTAGCAATAGGAACGATTATCGCTAATATAC
AGGTCATAAAAGCGGTTCATATTTTTGGTATTACGGCTACACTTGGAAATG

```
     TCATGTTTGCTTCTATATATTTAGCTACTGATATATTAAATGACATCTATG
     GTCGTAAAGTTGCTAAAAGAGCGGTGTGGCTTGGTTTCTCTTCTACCTTAG
     TAATGATTATAGTCATGCAAATGTCATTGCATTTTATTCCTGCCCCAGTAG
     ACAATGCGCAAAACTCATTAAAAATGATTTTTGATTTAGTGCCTAGAATTG
  5  CTATAGGTTCCATTATTGCTTATATCATAGGCCAACATATTGATGTATTTA
     TATTCAGTATGATTAAAAAGATATTTAGCTCTGATAAGACCTTTTTTATTA
     GAGCATATGGTAGTACCATTTTAAGTTCTATCATTGATACCGGTTTATTTG
     TTTCAATTGCTTTTATTGGTACTATGCCTGGTACTGCTGTTTTTGAAATAT
     TTATTACCACTTACTTGTTAAAACTAGTGTCAACTATTTTTAATGTACCAT
 10  TTGGATATATCGCTAAGTCACTATATCGAAAAGGAAAGATAGAACAACTAG
     ATAATGGGTATTGATATTTATTGTATCAGAAATATAGATATACAAAAGCGA
     ATCAGAGTTCCATGTGCTGAACTAGCTTCGAAATTCAATTTGGTTTGTCCG
     TAGAATTTTCGAGCAAAATTCTCAATGATAGGACCTACGACCAAAGTTAAA
     TAAATCTTAATATAAGGGCTAAATTTAGACATCAATGCCCTATTTGTTAGC
 15  AGAACCTGAAGTATATATGTTATCTTAGGTTCTGTTTGTTTTTGATTTAAT
     TTTTAGTGTTTTGTCTTTTTAGTTTGATACAATAGAATAGATACATTTTAG
     TGAGTAGACAATTCTGAAGAGGAGTAAAAGTAATGGCTAAAATACATTTTG
     ATGCTGCGACTAAAGGAAATCCCGGCCGAAGTGCTTGTGCGATTATTATTA
     AAGAAAATTCACAAAGATATACATTTACCCATGATTTAGGTGAAATGGATA
 20  ATCATAGTGCAGAATGGGCAGCAATGTTACACGCTTTGGAACATGCACGCG
     AATTAAAAGTATCTAACGCGTTACTTTTTACTGATTCAAAATTAATTGAAG
     ATAGTATGATGCAAGGTAAAGTTAAAAATGCTAAGTTTAAAGTTTATTTTG
     AAAACATAGAAATCTTAGAGCAAAGTTTTGATTTGATGTTTGTGAGATGGA
     TTCCACGAAAGCAAAATAAAGAAGCGAATCAACTTGCTCAACAAACACTAT
 25  ACAAACTTACATCATAATAATCTAAAGTGTTAATATGATATGTTACGTATT
     TAATTGTTTGTTCAACCTTACCAATAAGATTTCGGCTTTGAAAAGCTTATT
     ATAAATATGTTTGTACTGAATTATCGATGGTTAACAAAAAACAAGCCTGGGA
     CATTAAGTCCCGGGCTTGTTCTGTGTTAAAGATATATTTTACTTTTTAGAA
     CGCTTTTTGGTTTTTACAACTTTTTTAGTTGATTGAGGCGTTTTCTTCTTT
 30  TTAGAAGTAACTTTCTTCTGTTTTGTATCTTTGTTATCACTGTTTTGTTTA
     CGTTTAACGACGGGGATTGAATGAATATCTCCTTCACTCAATTGCTTTTCT
     TTAGATTTCTTATGTTTAGCAACTGGAATTGAATGTTTGTCATCGTTTTGA
     ATTTGTTCTTCTTTATCTATTTTACGCTTAGTTATTACAATTGGATCTATT
     TTATCAGTATCTGAAAGACGAATGTCTTTGTGATAATTTTTTATCGATTGT
 35  TTTTCTTCTTCTTTTTTACGTCTTTTAGCCAAAACTAACCAGAAACTAGAA
     AGTAAACCAGTTAATGTAATAACACCTACAGTTTTGCCGAAATTATCAACT
     AAATGTTTGAATAGG
```

```
 40  Sequence 3437
     step.1003b07.cons.ok
     CTTCAGCAGCAAAGTTAGCTTCACCTGAAATTCCATGTCTTTCATGTTGAA
     CTCTAAGCAACATAACAATATCTACTTTATCAATGACTTCATCAATTTCAA
     CATAAGGCGCCTCTAATGTATTATCTACCCATTCTTTTGGACTTGAGAACA
 45  TTACGTTGGCACCTAATGATGTTAAACTATGATAATTACTTCTTGCGACAC
     GAGAATTTTTAATGTCCCCACATATTAGAATATTCAAACCTTCAAACGAAC
     CATATTCTTCATATATTGTCATTATGTCTAATAAACTCTGAGTAGGATGTT
     GTCCACTTCCATCACCTGCATTAGCAATTGGAATATTTAATTGATCCAGTT
     CTTCGTAATAAGAATTTTGGGAGTGACGTATGACAAGTAAATCAACACCTA
 50  TACTTTCAAGTGTTTTACATGTGTCATAAAGTGACTCACCCTTTTTTACAG
     ATGATGTACTTGTTTCAAAATTAATAAGTTTTAATCCTAATTTTTGTTCTG
     CCATCTCAAAGCTACACTTTGTTCGCGTTGAATTTTCGAAGAATAAGTTTG
     ATACGTATTGACCGTTAAATTGAGGTAATGGTCGCTCACCAGATTTGAATT
     GGCAAGCGATAGTAATTAAATCATAAATTTCTGAATTAGATAAATGCTCCA
 55  TTGATAATAAGTGTTCCATAAAGCGCCTCCTATAGAAATAAAGCTTAATTT
     ATTTGGTTTTTATCTTTTGGCAAAATTAAATTTAATATTATTCCTGAAAGT
     GCTGATAATGCCATTCCTTCAATTTGTAAATTGATACCTATGCCTTTTAAA
     TTGATAAGAAGATTACCAATCCCGACAACAAGCACAACTGATGCTATAACC
     AAGTTGCGATTGCTTGCGAAATCTACTTGACTTTCAACAAGCATTCTTAAA
```

```
     CCACTAGCTGCTATAATACCGAATAATAAAATTGAGACACCACCCATCACT
     GGCGTAGGTATTGAAGATATTAAAGCAGTGAACTTACCAATAAATGCAAGA
     ATGATAGCTATAACTGCCGCACCACCAATAACGTAAATACTATATATTTTG
     GTGATCGCTAGTACACCTATATTTTCACCATAAGTTGTACTAGGAGGACCT
 5   CCTATCATACTTGCAAACATAGTTGAAACACCATCACCAATGATTGATTTA
     TCTAAACCTGGATTTTCAAAGAAATTGCGTCCTACTATTTTATTAATTACC
     ATTTGATGACCAATATGTTCACTTACCGTCACAAACACCACGGGTATCATC
     ACGAGAATGAGTCCTAAATGAAAGATGGTGTGTAATCTTTAAATGGTAGA
     TAAATATGAGGAAAATCTATCCATTTCGCTTGTGCTATTGGAGCAAATTTA
10   ACTATGCCCATGAAATGGATACAATATATCCCACTATAATACCTATAAGT
     ACAGGTATTAGTGATAAAAATCCTTTGAAGAATCCTTGGACGATGATGGTT
     ACTGCTAATGTAATCAAAGCAACAATTAAGTAACTTAAGTTATACCCTTTC
     ATTTCAGCAGAATTTTCGAACATGGCCATGTTTACTGCTGTAGGAGCTAAA
     CTTAACCCAATGACCATTATTACTGGTCCGACAACTACTGGTGGTAACAAG
15   TGCATTAACCAATGTGTTCCACTCAATTTAATGAAAATACCTATAATCACG
     TACATTAATCCACTCATAAATAATGCAACCAGCATATCTCCAAGACTATGT
     GTACTTAACCCTGTAATGATAGGTGTAATAAAGGCAAAACTCGATCCTAGG
     TAGGCAGGAATTTTCGCTTTTGTGATTAATATGTATAAAAGTGTACCAATA
     CCTGATGCCAACAAAGCTGCTGATATAGGCAGATGTGTTAAAAATGGTACA
20   AGTACAGTTGATCCAAACATTGCAAATAAATGTTGAAGACTCAGAAATGCC
     CATTGTGCTGGTTTTGGTTTATCTTTTACATCAAGAACTGGCTGAACGGTA
     CGCTCAAACATTTGCTCATTTTCCATGTTGATTCATTTCCTTTCATAAAAA
     AAGTCTCTTTACAATCAGTAGAATGTAAAGAGACTAAAATGAAATATATCT
     AAGTGTAATGTTAAGTTAAACTTTTATCATCACATTTATCAATGAAGCTTC
25   TATTGGATATTTATAGTTTACATAATCTTTTTATAGTCATATGTAGATTTA
     TCCATAAGAAACACTTCACCTTTTTCAGTCTCTCGTACTGTTTTAAAAGGA
     TAATTATTCAATGACGACTGCATTTCTATCGTCAATTTCTTCTAAATAAAC
     TGAAACAGATTCATCTCGTGCTGTAGGTATATTTTTTCCTACAAAATCTGC
     GCGTATAGGGAGTTCACGATGACCACGATCCACAAGTGCTGCAAGCCCTAT
30   TTTAATAGGTCTTGTATGTAATAAAATCGCATCTAATGAGGCTCTTACTGT
     ACGTCCGGTATACAAAACATCGTCAATGATAACAACCACTTTGTTATTAAT
     ATCTACATTAATATCAAAAGCGTATTGATCAGCTTGTTGCACTACCTTATC
     AACGTCATCTCGAAAATGCGTGATATCGATAGTACCTGTTGGTACTAATTG
     TTGTTCAATTGAATTTATTTTATCTTGTATACGATGTGCTAAAAAAGCACC
35   TCTTGTTTTAATGCCTAATAGAACTAAATCTTTAGTTCCCTTGTTATATTC
     TAGAATTTCATGAGCAATTCGTGTAATTGTACGTTGTATCGCTGCCTCATC
     TAAAATAATTCGTTCAGACATTGTTTCACCTCTTTTAATATAAAAAAAACTC
     ATATCTTTAATAAAGATATGAGACAAAAAGTTAAAGAGAATAATCTATATT
     CTCTATTATATTCACATTCTTAAGTTATAAACACAGCAAAAGTATTAGTAA
40   CTTAATCATCAAAATATTATTATCCTAATTATTCATATCTTCGGAGCCTCT
     CTGGACCCTCAATTAAAGATGACGTTTTTATATTTAAACTGTTATTAGTTT
     AAAACTTTTAATAGTATAAGTCAATGGCAAAGTTAGTATTTATAATTAAAA
     ATCGATGCCTCATTCGTTTATGCATCTCTACGTCTTATATCATCTAGTAAC
     TTTTCAAAATCGTTAGGTAATGTTGCATGCTTCTCAATGTATTCATGTGTT
45   ACAGGATGTTCGAAACCAATTATACCAGCATGTAAAGCTTGTCCATCGATA
     TCTAACGTTTTTTTAGGTCCATACTTCGGATCCCCTACAAGAGGATAACCA
     ATATATTTCATATGTACACGAATTTGATGTGTACGCCCTGTTTCTAATTGA
     CATTCTATCAATGTATAATCTTTAAAATGCTCTAATACATTAAAGTGTGTC
     ACAGCTTCTTTACCGTCATCAACCACAGCCATCGATTGTCTATCGTTTTTA
50   TTTCGACCAATTGGTGCGTCAATAGTGCCATAATCATGAGGGATATTACCA
     TGAACTAACGCCGTGTATTTTCGTTTAACAGTTTTAGACATAAGTTGTTCA
     ACTAAATGACGATGAGCAACATCATTTTTAGCAACCATTAATAAACCAGAC
     GTATCTTTATCTATTCTATGAACTATGCCGGGACGTATTTCGCCATTGATT
     CCTGATAAATCTTTAATTTGATACATTAATCCATTTACTAATGTTCCACTA
55   TAATGGCCTGGAGATGGATGTACAACCATACCTTTTGGTTTATACTTTAGA
     GCACAGTGGCG
```

Sequence 3438

step.1003b08.cons.ok
CTGTAATTATGAGTGAACGACTAGTTAAAGATGATGTATATACATCTATTC
ATATTGAAGAGTATGAGTCAGAAGCTCGTGATACGAAATTAGGACCTGAGG
AAATCACACGTGATATTCCAAACGTATCTGAAAGTGCACTTAAGAACTTAG
5   ATGATCGTGGTATCGTTTATGTTGGTGCAGAAGTTAAAGATGGTGACATCT
TAGTAGGTAAAGTAACGCCTAAAGGTGTAACTGAATTAACAGCTGAAGAGA
GATTATTACACGCTATCTTTGGTGAAAAAGCACGTGAAGTTCGTGATACAT
CATTACGTGTACCTCATGGTGCTGGCGGTATCGTTTTAGATGTTAAAGTCT
TTAATCGTGAAGAAGGAGACGACACTCTATCTCCTGGAGTTAACCAATTAG
10 TACGTGTATATATTGTCCAAAAACGTAAAATACATGTTGGGGACAAAATGT
GTGGTCGTCACGGTAACAAAGGTGTTATTTCTAAGATTGTTCCAGAAGAAG
ATATGCCTTATTTACCAGATGGTCGACCAATCGATATTATGTTAAACCCAT
TAGGTGTACCATCACGTATGAACATTGGACAAGTATTAGAATTACATTTAG
GTATGGCAGCTAAAAACTTAGGAATTCACGTTGCTTCCCCAGTGTTTGATG
15 GTGCCAACGATGATGACGTTTGGTCAACTATCGAAGAAGCTGGAATGGCGC
GTGACGGTAAAACTGTTCTATACGATGGACGAACTGGTGAACCATTCGATA
ATCGTATTTCAGTTGGTGTGCATGTACATGCTTAAACTTGCACACATGGTAG
ATGATAAGTTACATGCACGTTCTACTGGACCATACTCACTTGTTACACAAC
AACCATTAGGTGGTAAAGCACAATTCGGTGGTCAACGTTTCGGTGAGATGG
20 AGGTATGGGCACTTGAAGCTTATGGTGCTGCTTATACTTTACAAGAAATCT
TAACTTATAAATCTGATGATACTGTAGGTCGTGTGTTAAAACTTACGAGTCTA
TAGTTAAAGGTGAGAACATTTCTAGACCTAGTGTTCCAGAATCATTCCGAG
TATTGATGAAAGAACTACAAAGTTTAGGTCTAGATGTTAAAGTGATGGATG
AGCATGATAATGAAATCGAAATGGCTGATGTTGATGACGAAGATGCAGCAG
25 AACGCAAAGTAGTAAAATATACTGGATTAAATAACATAATTTGATGAATCA
AACCACTTACGCCGTGATTTTTTGGCACCCATAAAATTGGAGACATATAAA
ACAATATTCTCATTAATGCTTGCATAATCATCTGCGTATCTCTAATTAAAA
TCCCCAAAGTGGATGTTAAAAGTGCCACCGACGATGTTAGCAAATATGCAA
AAGGTACATATATAAGTAATTGTACAATGTGAATTGAAGGGATAATTCCAT
30 TGAACATACAAGCTATTATAATAATTGCTAATAATCCTAAATGACCATAGA
ACCTACTTGTTACAATATAAGTAGGAATGATTGAGAGTGGGAAATTCATCT
TTGCCACTTGATTGAATTTCTGTGAGATTGATTTAGTTCCTTCTAGGACAC
CTTGATTAATAAAGAACCACATACTTATACCAACTAACAACCAAAATACAA
ACGGAATGCCGTGTATAGGTGCATTGCTTCTAATTCCTAAACCAAAAACAA
35 GCCAATACACCATGATTTGTAAAGCAGGATTCAACACTTCCCATGCAATAC
CTAGATAGTTATTATGATTGGTAATTTTAATTTGGAATTGTGCTAGTCTTT
GAATTAAATAGAAGTTTTTAATATGTTCTTTTAAAACCGTTCCAACTGCGA
ACATTAAACTTAACCACGCTTTCTCCTATAATAGATTATCTATCTATTGTA
AATCATTAATAGTTTCTCTTTAATTTGATTCATTGAGTCCATTCTTAAATA
40 ATAACATTTAACTGTTTATCATAAGCTCATTATATTATACAATCAATTCAG
TTAAAAATTGTTAGAAATTTCGTGTAAGATTTAATTTTTAGAAAATGAATCA
CTTTCTTAACTCACTGTGTTATATTTATTTTCAACAAAATCTTCTATTCAC
TTTACACTAGGTTAATAAATTTAACAACTGATGTGTTGTATATTATAATCT
ATATAAATGTTTATAAGGAAGGTACAAAAATGAGCGTTTCGGTAAATATTG
45 AAAATTTAACAAAAGAATACCGTATCTATAGAAATAATAAAGATAGAATTA
AAGATGCATTAATACCTAAAAACAAAAATAAAACATTTTACGCTCTGGATA
ACGTAAGTTTAACAGCGCATGAGGGAGATGTGATAGGTTTAGTCGGCATCA
ACGGTTCAGGTAAGTCTACTTTAAGTAATATGATTGGTGGCTCTATTTCAC
CAAGTTCCGGTGAAATAACGAGACATGGTGATGTGAGTGTCATCGCTATTA
50 ATGCAGGACTAAATGGACAATTGACAGGTGTAGAAAATATTGAATTTAAAA
TGCTCTGCATGGGCTTTAAAAGGAAAGAAATTAAAAAATTAATGCCGGAAA
TTATAGAATTTAGTGAACTCGGCGAATTTATTTATCAACCTGTTAAAAAAT
ATTCAAGTGGTATGCGTGCAAAACTTGGATTTTCAATTAATATTACTGTTA
ATCCTGACATATTAGTTATTGACGAAGCATTATCAGTAGGCGATCAAACAT
55 ·TTACTCAAAAATGTTTAGATAAAATTTATGAATTTAAAGCGGCTAAAAAAA
CAATATTTTTTGTTAGTCATAATATTAGACAAGTGCGTGAATTTTGTACAA
AAATCGCTTGGATTGAGGGCGGTAAACTAAAAGAATTCGGCGAACTTGAAG
AAGTATTACCTGATTATGAGGCGTTTCTTAAAACTTTTAAGAAAAAATCTA
AAGCAGAACAAAAGGAATTTAGAAATAAATTAGATGAGTCACGTTTTGTCG

TAAAATAAATTTGAAGAAAGTAAAACTTAAACATTAAGGGAGTGCAATAAA
ATTCATCCGAATTGATTGCCTCCCTTTTTTTACTCTTATTTTGATTTCATTT
TATATTTTTGTATTGCTACCTTTAACATAAATTTAGGAATACTTATCATTC
TCCCAATGCGTTTCCAATCAATAAGCACACGATATACCCACTCAATATTTA
5   ACTTTCTAAATATTTGAGGTGCTCTTTTCTTTGAGCCACTGAATACTTCAA
ACGACCCACCTACGCCCATCATCACAGTGTGCTTAAACTTGTCCTTATGCT
TTTGAATCCATTGTTCTTGCTTTGGAAATCCCATTCCTACAAAAATGTAAT
CGGGATTAAAACTTGTTATACGTTTTATGACTGTTTCATCTTCTAGATGAA
TATAGCCATGATGATGTGCAAAGTGGATATTAGGGTATTGAGAT
10

Sequence 3439
step.1003b09.cons.ok
AGCAGAAGTAGCTGAAAACAACATTGTGAAGGAAAACCCAGATGTAGAAAT
15   CTTTGAAGAAGGTATTATTGATTCATTTCAAACTGTAGGTTTACTTTTAGA
AATTCAAAATAAATTGGATATTGAAGTGTCAATTATGGACTTTGATAGAGA
TGAGTGGGCAACTCCTAATAAAATTGTAGAAGCTTTAGAAGAGTTACGATG
AAACTGAAACCATTCTTACCCATTATAATCAGTGGCGTCATATTTATTATA
TTTTTATTTTTACCAGCAAGTTGGTTTACTGGTCTTGTGACAAATAAGACG
20   TTAGCTGATAATAGAATTTCTTTAACTGATCAAGTTTTAAAAGGCACGCTC
ATTCAAAATAAATTATTCGAGTCAAATAAGTATTACCCGATTTATGGTTCA
AGTGAACTGGAAAAGGATGATCCTTTTAATCCAGGTATAGCGCTAAATAAA
CAAAATGCTAGTAAAGAAACGTTCCTTATTGGTGCGGGAGGTTCGACAGAT
TTAATTAATGCAGTTGAACTTGCAGCGCAATACGACAATTTAAAAGGAAAG
25   AAATTTACATTTATTATTTCTCCACAGTGGTTTACGAATCATGGATTAACA
AATCAAAACTTTGATGCTCGTATGTCACAAAGTCAAATCAATCAGATGTTT
AATCAAAAAGATATGCCAGCTAATTTGAAGAAACGGTATGCACAAAGATTG
TTACAGTTTCCGCATGCACACAATAAGTCATACCTTAGAGAACAAGCAAAA
CATCCTAATGATGTCTCTGGAAACTACATTTCTTCATTTAAAGAAAATCAA
30   TTAACTAAGATTGAAGCTATTAAATCATTATTCTCATTCACTAAGCCACCT
CTAGCAGAAGTAAAACCTGCAACAAGAAGAAGATGCTTCATGGGATGAGATG
AAACATAAAGCTGCCGATATAGGCAAAGCAAATACTCAATCTAATAAATAT
GATATAAGAGATCCATATTGGAAATTGATAAAACAAAACAAGCGTAAAATC
AAAAGGGATTATGAGTTCAACATTAACTCACCCGAGTTCCAAGATTTAAAA
35   TTATTAGTGCAAACGCTACATGCTGCTGGAGCTGATGTACAATATGTTTGT
ATACCTTCAAATGGAAGATGGTATGATCATATAGGTATCAAAAAAGATAGA
CGTGAAGCTGTATATAAAAAGATTCACTCAACTGTAGTTGATAATGGTGGG
AAAATTTATGATTTGACAAATAAGGACTATGAAAAGTACGTAATTAGTGAT
GCTGTTCATATTGGATGGAAAGGTTGGGTTTACGTCGACCAGCAAATTGCA
40   AGACATATGGATGGTCATGCGCCTAAAAATCATGAAGTCGATTATTCTAAA
AATAAACCACCGCACAAACATCACAACGATCGTCAAGATGATCAACATCAA
GGCAACAAATAATATTATTAACCGAAATAAAATTAAGTGAGCTAGGTACAG
AATTCGTTATTTCATCCCGTGTAACGAGATTTATAAAGGAAATACCTCGAC
GTAATGTATTTTTCGTTTACTTCTTTGATAGCCAGTATATAAAATGCTTAG
45   GAACAATGTTTTGTATCCTAAGCATTTTATTGTTTTATTTAGCATCATAGA
GTTAATATTAGTGCGCTATGACTTCATTTATTAATTTCTATCAAATTAAAA
AACCTAGGACGTCACTATCTATCCTAGGTTAATAAATAATTATATAGGGAT
GTATTAGAATACTTGTTCTACTTCTATTACACCTGGAACTTCTTCATGAAG
TGCACGCTCGATGCCAGCTTTTAAGGTAATTGTAGAACTAGGACATGTTCC
50   GCAAGCGCCATGTAATTGTAGCTTAACAATACCATCTTCAACATCTACAAG
CGTACAATCGCCACCATCTCGTAATAAGAATGGACGCAAACGTTCTATAAC
TTCTGCTACTTGATCAAACATTGTTGGATTCTCAGTAGGCATGTGACATGT
CTCCTTTCAGAGCAATTATTTCATAATCATTATTATTACATTATAATAGAT
AAAGTAGTAAAAATCTATAAAACAAGATAGGGGATTTTAATGACTAAAATT
55   AGTGTTGTCGTATATGGAGCAGAAGTCGTTTGTGCGAGTTGTGTAAATGCA
CCTACATCTATAGATACTTATCAATGGCTTCAAGCATTACTTTTAAGAAAG
TTTCCTCAACATCATTTTGAATTTACATATATTGACATACGAAATGATACT
GAAAATTTAACTGATCATGATATGCAATTTATAGAAAGAATTAATGAAGAT
GAATTGTTTTACCCATTAGTTACGATGAATGATGAATATGTAGCAGATGGT

TACATACAATATAAACAAATAACCCGTTTTATTAAATCATATTTTACTATG
TAAGATTAAAAGAGGGGGCAATATTCATACTAACGATGTGCCTCAATTGAC
AGTGTCTGTGTTATTTTTTAGTGAAAATCTTCATTTTTAGGCTATATCTTT
TATTGAGAAGAAGGACATTTCAATGCTGTCATAGTTTGTGATTATTAATTA
5   AGCACCTAGGGAAATGTAAATGCCCTAGGTGCTTTATTTATTAACCATTAT
GATATTTATAGCGCCACAGAACACCTGATTTTAGAATAGAGGCTAATCGTC
CGGTAACTGTTCGATCCATGATATAAGCAAAACCTTGTTTGTCACCTAAAG
AGCCTAAAAAGCCTTGTACTTTAATTTCAGGCATTTTATCTGGAAGTGGTT
CGTTATTCCATTGCTTCTTCAACACCTCAGCAATCTGTTCACCTTGTAGTT
10   CTGCTAGTTGAGCACTGGGTGCATGTGGTAAATTAGCACAGTCACCTACGA
CATAAACATTTCTATAGGTTGGGACTTGATGGTACTGATTAATAATTACGC
GTCCAGTGGTACTCATATCAATAGGAAGATTACGCACAATTTCAACAGGTT
GTATGCCTGCTGTCCAAACGACTAAATCAATATTTTCTGGTTTACCATTAT
TATAAATTTTTCCGGGTTCTACTCTGTCGATGACTGAATTAGGTACTACAG
15   TAACATTGTGTTTAGAAAACCAATTAGATATGTATTTACTCAGTTTCTCTG
GAAAATTCCTTAAAATTCGAGGCCCTCTATCATATAACAAAATTTCCAAGT
CTGATCGACTTTCACGTAGCTCGCTCGCTAATTCAATGCCACTTAAACCTG
CCCCAACGATACCTACACGTGCCCCTTTAGGTAACTCGCTAATTCTATGGT
ATGTTTCACGCGATTTAGATAATGTTTGAATGCTATGTGTATATGCTTCAG
20   CACCAGGGACATTATGATATTTATCCTCACACCCTAGACCAATGATAAGTT
CGTCATAATCTATTTTTGAATTTCCAACTGTTATCATTTGTTCGTCCAAAT
CTATATCACTGATTTCCCCATAAACCGTATTAATTTGACTGCTATCTGGAA
ATTGGATTCGCACCTCTTTGTCAGATTTAGTTCCTGCTGCAAGTGCATAAA
ATTCAGGTTTTAAACCGTGGAATGGCATGCGGTCGATTAAAGTTAAGTGAT
25   ATCCCTCAGGAATTGAATGAGGTAAAATGCGCGACATAATTCGCATATTAC
CATAGCCCCCGCCTAGTAATACTAAGTTTTTCATTGATTCTTTAGAGCACA
GTGGCGATGATATCAGAT

30   Sequence 3440
      step.1003b10.cons.ok
      ATTCAACCACTTTACGAATATATATAATAATAACGTTTGCATTACCATATT
      AACTATCTGAGTGATAGGAAATCTTAAAAACTTTTCTATTGTAGGTTTTAC
      TTTATATACAAAATAACAATTTAAATAATATGAAATAATAAAACTTACAAT
35   AAATCCAACAATGTGACTAACCATATAATTCACATGTAACACCTTTAACAA
      AAATAAGTATGTTATATAGTAGTTAAAGGTATTAATTCCACCAACTATGAT
      AAACTTTATAATCTCATAATGTATTCGGGTTAACTTCATAATATTTCCCCT
      TTGAACAGTTATTGTTTTGTACCATATGGGACTACTTTAAATGTGATGTCT
      TTGGGCAAATCATAATAGTGTTTGAAATCTTGAAGTTGTTCCTTCGACTTT
40   GGTGTAGTCATATGAGTATATCGCTCAAATGGTAATCTTTCCAATGTTATT
      TTCTGATAGCGATGATGTTGTGTTATTTGTTCTTTAATGAAGTCTATTCTG
      TGCACACTACTAATATGTATAAAAGTAAATCCAATCATCATAATGATGCTG
      ATGATAATCGCAAATATTTTAATTATATGTTCAAGTTGTTTAAATAGCACA
      TCACATTGCTGAATTAAACAAAGTAATATCACGATCCATAAAGTATAAATA
45   AAATAAAAATTTCTATAACTTATAGGCGTCACAAATAAAAGTGGCAAAACA
      GATGAAGCCATAGCTATAAAACTCCCCATCACAATCATTCTTATGTATCTT
      TGCTGTATCATTTTAAACACAACGTATATCACACTTATCATGTAAATGAAG
      CAAATCGTTGTATTCAAAACGGCTATAGAAAATGAAGCTTTATATAATTCA
      AAATGAAATTGATTGTAAACGAAGATCTTATAAATAGGTAAAGTAATTAAA
50   CCTAAGAGTAGTGGTATTTTAATATAAACTCTCATATGCTTCAGGCTTTTA
      TTTTGCTTAAGTAAAACTATACTTACTATTGATATCACGGTAAGAATAATC
      ATTTGATTAATAAACATATATTCTGGTACAAGCTTAAATAACGTCACACCT
      GCTTTATGTATCATTCCATGACTATCGGAAATTGAATAATGCGTATTTAAT
      CCATCCTTAATTAAAAAATAATTGAAGTTTAAAAACATTATAATGTTACCT
55   ATACAACTAAGCATAAATCCTACAATTAAGAAATAACTGAGTCTTTTTTTA
      ACAAAGAAATAGACTACCATTCCTATTAAAATAATTAAGCTATTAGCGATG
      GAAAGATTCTCCAAGAAGAATTGTCCAAACAAACTTACTAATAAAAATACC
      CATAATTGCATTTCAGAAACTGTATCGTGCGACTCAATCTTTTTAACTACC
      GTAAAAAGAATAAAAAGTGATAGGACTGTAGCAGGTATATAACTAAAAAAT

CCAGTAAACCACCCGTAAGTTTCGCTATAAATTGTATTAGGTACAGTAACC
ATTAACACAAAACTCAAAATAAAATAAATACGATTCGTATGTAATTGAACC
ATATAAGCAACTAAATAGATAACTAAAAACGAAGTTATCGCATATATTAAA
GCTCTTATTATGTTATTATGTACGGCAACCCATTCAAATAAATGTCCTAAA
5   TAACGACCATTTTCTTGAGTCAAATATTGACTTAGATTAACTTTATATGCA
TGCCAATCCGTAGTTGATAATGGTGTTAGAATTGCCATTATAAGATAAAAG
ATAAAAATTCCTATTAACCAAAAAAATGATTCATATTGTTTTAAGTTATTC
ATTGTTGACCTCTCTGTTCTTCATGATGTAATTCAAAACTAACATTTAAAC
GATTATCATGTTTATTCTCTTTTTTATGTGATTTATTCAATTATCTACTGA
10 AAATATATAGACATTTATTGAAGCTTTAATCATATTAATATATACATATTG
CTCATCGTCAATTTTATATTCTTATTATTATACAATTCACGTATGATTTCC
ATTTTAATTATCAAGATTTCTTTAGTTTAATTTCATTTTTGGATAGAATAA
ATATAAACCAATTGTTTAGGAAGTGATATAAATGTCACCATATAAAGTTAT
CATTGCACCTGATTCTTTTAAAGAAAGTATGTCGGCGAAGGAAGCTGCTTT
15 AGCTATTAAAGATGGATTCCAAGAGGTGTTCGATTCCAGTACAATATATGA
CATTATTCCTATGTTGAAATAAGTCATCGTATTCATGAACGCCCTGAATTA
GGCAATGAAGAAATTTTTGCATCGAGAACATTAATTGACCAATTAAGAGCA
AATCGATTCGAAATCGAAACGGATATTGCAGGACATGCAACAGGATTTATA
GCAACGTATGATTCTGATATGACTGGACCGGTTATAGGATTCTAGCTGAA
20 TATGATGCTTTACCTGGTCTTGGTCACGCATGCGGGCATAATATTATTGGT
ACTGCTAGCGTACTTGCTGCAGTAGCACTAAAAGAAGTCGTCGATGAAATT
GGTGGTAAAGTAGTCGTTTTGGGATGTCCTGCTGAAGAAGGTGGGGAAAAT
GGCTCCGCAAAAGCTTCTTATGTTAAAGCAGGTGTCATTGATGAAATTGAT
GTAGCATTGATGATTCATCCTGGAAATGAAACTTATCGTACAATTAATACT
25 TTAGCTGTGGATGTTCTAGATATTAAATTCTATGGACGTAGTGCGCATGCA
TCTGAAAATGCAGATGAAGCATTAAACGCTTTAGATGCAATGATTTCATAT
ATTAATGGTATAGCACAGTTAAGGCAACACATTAAAAAAGGACAACGAGTT
CACGGGGTTATTTTAGACGGTGGTAAAGCGGCTAATATTATACCTGATTTT

30

Sequence 3441
step.1003c04.cons.ok
GTATAAGTAGCTAATTTTTCATCGTATAAAGTGTAAGGAGATTGTCTGCCA
35 TTGACAATAGCGTTCCCTTTAAATAATTTAATTCTCACATCTCCCTCAACA
TATTGTTGAGTACTATCGATAAAGAGTTTTAAACTATCTGTTAATGGCGAG
AACCACAAACCATTGTATATTTGTTCTGAAAATTGTTTTTCAATGACAGGC
TTAAAGTGCGCTACGTCTTTAGTTAATGTAATTGTTTCTAGTGCTTTGTGT
GCTTTTAAAATAACTTCCGCACCAGGTGTTTCATAAATCTCTCTCGATTTT
40 ATCCCGACCATTCTGTTTTCAACATGATCGATTCTACCAATACCGTGTTTG
CCAGCAAGTTGATTCAAGTAAAGAATAAGGTCATCTAATTGATAATCTTTG
CCATCAACTTGTACTGGAATACCTTGTTTAAATGTAAGGATAATTTCGTCT
GCATTGTCTGGAGTTTCTTCTAAAGGTGTAGTTAAATCAAATGCATCTTCC
GGAGGTGCGGCATACGGATCTTCTAAAATACCACATTCATTAGCTCTCCCC
45 CATAAGTTTTGGTCAATTGAGTATGGCGAGTCATAATTGATTGAAACAGGA
ATATTATGTTTGATTGCGTAATCAATTTCTTCTTCTCTGCTCCAAGCCCAT
TCACGAACAGGTGCAAATGCTTTTAACTTAGGATTTAAAGCTTTGATTGCC
ACTTCGAAACGTACTTGATCATTACCTTTACCAGTACATCCATGCGCAATA
CCAATAGAATTTGTTTTTTTCAGCAATTTCAACCAGTTTTTTCGCGATGAGT
50 GGACGTGATAATGCTGAAACTAGAGGATATGCATTTTCATACATTAAATTT
CCTTTAATAGCATAACTTACATAATCATCACTAAATTCTTTAGTTGCATCA
ATAATATGACATTCGACTGCACCCATATCTAAAGCTTTTTGATATACAACG
TCTAAATCTTTGCCTTCGCCTACGTCAAGACAACAAGCAACTACATCATAT
CCTTTATCAATAAGCCATTGAACTGCAACGCTTGTATCTAAACCACCTGAA
55 TATGCTAAAACGATTTTATCTTTCATATTTACACCTCAAATAATTTTTTAT
GATTATTAATATGCATCTATCATATCAGATTGATAATGAAAATAAAACCAT
TATTTTACATAAATATTCATTTTTAAATCTGATTATATTTAGATATTTTTG
TTTGAAATACTGTTATAAAAGCTTTTATATATATTTTAAAATTTCTGTAAA
TTGTATATCTAAAGTGGTGGTGTGACCTTTGATATTAAAAATAACTTTTTT

```
   ATGAATGAAAGGTTACATAAAAAATAATTTTCCTGTGATAAGTTAGCCTTG
   AGAATTGAATAGCGCATCTAAACTTAGTAGAATTAGTATAAGTACTTATAA
   CAGGAGGCAATAAGATGACTCACATTCAATTAGACTATGGCAAAACTTTAG
   AATTTTTTGATAAGCATGAACTAGATCAGCAAAAGGATATTGTTAAAACTA
 5 TCCATCAAACTATTCATAAAGGTACAGGAGCAGGTAATGACTTTTTAGGTT
   GGTTAGATTTACCTGTTGATTATGATAAAGAAGAATTTTCTAGAATCGTCG
   AAGCATCTAAACGTATCAAATCAAATTCCGATGTACTTGTTGTTATCGGTA
   TTGGAGGTTCATACTTAGGTGCACGTGCTGCAATCGAGATGCTTACATCTT
   CATTTAGAACAAATACGGAATACCCTGAAATTGTATTTGTAGGTAATCATT
10 TATCCTCAAGTTATACAAAAGAATTACTTGATTATTTACAAGGAAAAGATT
   TTTCAGTTAACGTTATTTCAAAATCAGGTACTACGACAGAACCAGCAGTTG
   CATTTAGATTATTTAAACAATTGGTTGAAGAAAAATATGGAAAAGATGAAG
   CTAAGAAACGTATTTTTGCAACGACAGATAAATCTAAAGGTGCACTTAAAC
   AATTAGCAGACAATGAGGGTTATGAGACGTTTGTTGTACCTGATGATGTGG
15 GAGGTCGTTATTCTGTTCTTACAGCTGTAGGATTACTACCAATTGCAACTG
   CAGGTATCAATATTGAATCAATCATGATTGGTGCGGCTAAGGCACGTGAAG
   AGTTATCTTCTGATGATTTAGATCAAAATATCGCATATCAATATGCAACTA
   TTCGAAATATTTTATACAGCAAAGGTTATACTACTGAAATGTTAATTAATT
   ACGAACCCTCTATGCAGTATTTCAACGAATGGTGGAAACAATTATACGGTG
20 AATCAGAAGGGAAAGATTTCAAAGGTATTTATCCATCAAGTGCGAATTACA
   CAACTGATTTACATTCCTTAGGACAATATGTTCAAGAGGGCCGTCGTTTCT
   TATTCGAGACAGTGGTTAAGGTCAACCATCCAAAACATGATATCAAAATTG
   AAGAGGATGCAGATGATTTAGACGGACTGAACTATCTTGCTGGCAAATCAA
   TCGATGAAGTGAATACTAAAGCATTTGAAGGTACATTACTTGCACATACCG
25 ATGGTGGCGTTCCAAATATCGTTGTAAATATTCCTCAGTTAGATGAAGAAA
   CATTTGGATATGTTGTTTATTTCTTTGAATTAGCTTGTGCAATGAGTGGAT
   ATCAATTAGGTGTTAATCCATTTAATCAACCTGGAGTTGAAGCCTATAAAC
   AAAATATGTTTGCGCTATTAGGTAAACCAGGCTTTGAAGATAAGAAAAAAG
   AATTAGAAAATCGTTTATAATTTTAAGTTGAACTTAAAAATGAGCCTAGGA
30 TTATCTAGGCTCATTTTTATTATGTGTATTAATAAGAGTAAATGAAAATTG
   ATTAACGTCAGATTAATAAGAAGTCAAATTCAGATAGGAAATATTAAATTT
   TCAATTCATTAGAGTTTCGATAAAGAATGGAATAATAAAATTGATGAATTA
   GATATCAATATATAGATTTATTAATTCCTACCTAAGTCTGATTTATGTTGT
   TGACCTAATTTTAATGTTTCAAAAAATATCATTTACAACTATAATGAATAA
35 GTAATACAACCCTAAGAAAGGATTTGATGTATTGTCGCTTTCTCAGTTAGA
   AGAATGGTTTGACGCATTTCGACAATTCGGATATATTCCTGG
```

```
   Sequence 3442
40 step.1003c05.cons.ok
   ACATTCAAAACTAGATAGTAAGTAAGATTTTGCGTCGCAAAACGTTTTTTA
   AAAATTGATTAAGTCTTCGATCGATTAGTATTCGTCAGCTCCACGTGTCAC
   CACGCTTCCACCTCGAACCTATTAACCTCGTCATCTTCGAGGGATCTTATA
   ACCGAAGTTGGGAAATCTCATCTTGAGGGGGGGCTTCATGCTTAGATGCTTT
45 CAGCACTTATCCCGTCCATACATAGCTACCCAGCTATGCCGTTGGCACGAC
   AACTGGTACACCAGAGGTATGTCCATCCCGGTCCTCTCGTACTAAGGACAG
   CTCCTCTCAAATTTCCTACGCCCACGACGGATAGGGACCGAACTGTCTCAC
   GACGTTCTGAACCCAGCTCGCGTACCGCTTTAATGGGCGAACAGCCCAACC
   CTTGGGACCGACTACAGCCCCAGGATGCGATGAGCCGACATCGAGGTGCCA
50 AACCTCCCCGTCGATGTGAACTCTTGGGGGAGATAAGCCTGTTATCCCCGG
   GGTAGCTTTTATCCGTTGAGCGATGGCCCTTCCATGCGGAACCACCGGATC
   ACTAAGTCCGTCTTTCGACCCTGCTCGACTTGTAGGTCTCGCAGTCAAGCT
   CCCTTATGCCTTTACACTCTATGAATGATTTCCAACCATTCTGAGGGAACC
   TTTGAGCGCCTCCGTTACCTTTTAGGAGGCGACCGCCCAGTCAAACTGCC
55 CGCCTGACACTGTCTCCCACCACGATAAGTGGTGCGGGTTAGAAAGCCAAC
   ACAGCTAGGGTAGTATCCCACCAACGCCTCCACGTAAGCTAGCGCTCACGT
   TTCAAAGGCTCCTACCTATCCTGTACAAGCTGTGCCGAATTTCAATATCAG
   GCTACAGTAAAGCTCCACGGGGTCTTTCCGTCCTGTCGCGGGTAACCTGCA
   TCTTCACAGGTACTATGATTTCACCGAGTCTCTCGTTGAGACAGTGCCCAA
```

ATCGTTACGCCTTTCGTGCGGGTCGGAACTTACCCGACAAGGAATTTCGCT
ACCTTAGGACCGTTATAGTTACGGCCGCCGTTTACTGGGGCTTTGATTCGT
AGCTTCGCAGAAGCTAACCACTCCTCTTAACCTTCCAGCACCGGGCAGGCG
TCAGCCCCTATACATCACCTTACGGTTTAGCAGAGACCTGTGTTTTTGATA
5   AACAGTCGCTTGGGCCTATTCACTGCGGCTCTTCTGGGCGTGAACCCTAAA
GAGCACCCCTTCTCCCGAAGTTACGGGGTCATTTTGCCGAGTTCCTTAACG
AGAGTTCGCTCGCTCACCTTAGAATTCTCATCTTGACTACCTGTGTCGGTT
TGCGGTACGGGCACCTGTTATCTATCTAGAGGCTTTTCTCGGCAGTGTGAA
ATCAACGACTCGAGGAAACAATTTCCTCTCCCCATCACAGCTCAGCCTTAT
10 GAGTGCCGGATTTGCCTAACACTCAGCCTTACTGCTTGGACGTGCACTCCA
ACAGCACGCTTCGCCTATCCTACTGCGTCCCCCCATCGATTAAAACGATAC
TAGGTGGTACAGGAATATCAACCTGTTATCCATCGCCTACGCCTGTCGGCC
TCAGCTTAGGACCCGACTAACCCAGAGCGGACGAGCCTTCCTCTGGAAACC
TTAGTCAATCGGTGGACGGGATTCTCACCCGTCTTTCGCTACTCACACCGG
15 CATTCTCACTTCTAAGCGCTCCACATGTCCTTGCGATCATGCTTCGACACC
CTTAGAACGCTCTCCTACCATTGTCCAAAGGACAATCCACAGCTTCGGTAA
TATGTTTAGCCCCGGTACATTTTCGGCGCAGTGTCACTCGACTAGTGAGCT
ATTACGCACTCTTTAAATGATGGCTGCTTCTAAGCCAACATCCTAGTTGTC
TGGGCAACGCCACATCCTTTTCCACTTAACATATATTTTGGGACCTTAGCT
20 GGTGGTCTGGGCTGTTTCCCTTTCGAACACGGACCTTATCACCCATGTTCT
GACTCCCAAGTTAAATTAATTGGCATTCGGAGTTTGTCTGAATTCGGTAAC
CCGAGAGGGGCCCCTCGTCCAAACAGTGCTCTACCTCCAATAATCATCACT
TGAGGCTAGCCCTAAAGCTATTTCGGAGAGAACCAGCTATCTCCAAGTTCG
ATTGGAATTTCTCCGCTACCCTCAGTTCATCCGCTCACTTTTCAACGTAAG
25 TCGGTTCGGTCCTCCATTCAGTGTTACCTGAACTTCAACCTGACCAAGGGT
AGATCACCTGGTTTCGGGTCTACGACCAAATACTCAACGCCCTATTCAGAC
TCGCTTTCGCTGCGGCTCCACATTTGCTGCTTAACCTTGCATCAGATCGTA
ACTCGCCGGTTCATTCTACAAAAGGCACGCCATCACCCATTAACGGGCTCT
GACTACTTGTAAGCACACGGTTTCAAGTTCTCTTTCACTCCCCTTCCGGGG
30 TACTTTTCACCTTTCCCTCACGGTACTGGTTCACTATCGGTCACTAGAGAG
TATTTAGCCTTAGGAGATGGTCCTCCCAGATTCCGACGGAATTTCACGTGC
TCCGTCGTACTCAGGATCCACTCAAGAGAGAATATGTTTTCGACTACAGGA
TTATTACCTTCTTTGATTCATCTTTCCAGATGATTCGTCTAACATGTTCTT
TTGTAACTCCGTATAGAGTGTCCTACAACCCCAACAAGCAAGCTTGTTGGT
35 TTGGGCTCTTCCCGTTTCGCTCGCCGCTACTCAGGGAATCGATTTTTCTTT
CTCTTCCTCCGGGTACTAAGATGTTTCAGTTCTCCGGGTCTGCCTTCTGAC
ATGCTATAAATTCACATGTCGATAACATGACATAACTCATGCTGGGTTCCC
CCATTCGGAAATCTCTGGATCAACGCTTACTTACAGCTACCCAAAGCATAT
CGTCGTTAGTAACGTCCTTCATCGGCTTCTAGTGCCAAGGCATCCACCGTG
40 CGCCCTTAATAACTTAATCTATGTTTCCACCATATTTTGAATTGTTATTCA
AAATAAATAGCTAAAACTAGTTATTAATCTTGTGAGTGTTCTTTCGAACAC
TAGCGATTATTTATGAATTCAAGCTTATTTAAAACTCTATTCACTCGGTTT
TGCTTGGTAAAATCTTACTTGCTTATCTAGTTTTCAATGTACAAATGAATG
TTAATAAACATTCAAAACTGAATACAATATGTCACGTTATTCCCTCATCTT
45 CGTAGAAGATGTTCCGAATATATCCTTAGAAAGGAGGTGATCCAGCCGCAC
CTTCCGATACGGCTACCTTGTTACGACTTCACCCCAATCATTTGTCCCACC
TTCGACGGCTAGCTCCAAATGGTTACTCCACCGGCTTCGGGTGTTACAAAC
TCTCGTGGTGTGACGGGCGGTGTGTACAAGACCCGGGAACGTATTCACCGT
AGCATGCTGATCTACGATTACTAGCGATTTCACTAATTTCTTGAAAAGTTT
50 TTTCTAATTTTTCAAAGAAACCACCTTCTTGATGTTGCTCGTTTGAGGTTT
TTTGATTGGGTCTTTGATTTTGTCTGTCCTCGTGACTTTTATCTCGTAATG
CGCTAACACCTGTAATTATCACAGATATGATGAAAATGATAATGCCAATAT
TCATCTTATCACCTCTTCATTATTGTTGTGGTGATTCATCATCGTTTTGAT
CAGTGCGCTTATTTATTGCATTTCTCATACCTGTATCCGCTTCA
55

Sequence 3443
step.1003c07.cons.ok
TGTAGCAATACCACCATGCAACTGACCTACGTTTTCAAATCCTTCTTTTAC

```
TAACCAACCAGAAAATTTTTCACAACGAATGCCACCTGTACAATATGTGAC
AATATTTTTTCCGTCGAGTTGTTCTTTATTATTACGCACCCATTCAGGTAA
GTCACGGAATCGTGTTATATCAGGACGAATAGCTCCACGGAAATGTCCTAA
ATCGTATTCATAATCATTTCGAGCATCTAATATAACAGTATTTTCATCTTC
```
5
```
TAGTGCGGCTTTAAATTCTTTAGGAGAATAGTATTTACCGGTAATTTCACG
TGGATTAATATCTTCTTCTAAGTCAAGTGCAACAATTTCACGTCTTGGACG
CACGTGCATCTTTTTAAACGCATGACTTTCAGCTTCATCAATTTTAAAAGT
TAAATCAGCAAAACGACTATCTGCATGCATATGCTCTATATATTTATCAGT
ATCTTCTTTTGTTCCAGATAATGTTCCATTAATGCCTTCCGTTGAAACTAG
```
10
```
TATTCTTCCTTTTAAATGATGTTCCTTACAAAATTTCAAATGTTCGGCTGC
AAAAGTTTCAGGGTCATCTATAGTTACATATTTATAATAAAGTAGTACTCT
ATAATCCATAAATCTGTGCTCCTCTTTAGTCAATTTTAGTTAATTTAAAAG
TATTTGAAGTTATATAAAAAAATGATAGTGATTGAAGTTTAAAAATTAGAC
ATTTTAATATTGTATCAAAATGAGAATTGCTTTCAATAAGATAGCCTATAA
```
15
```
ATCCGTCATTTGACCGATATAAGATGTGGTGGTTTTCATATATGAATAAAA
GTGTAGTCGTGTAGCTATAATGACATTTAATATAAATTTTTATAAAATGAG
AATAAAGTCACTGAATAAAGAGAGGGATTCAATCGATGGCAAATCAAAAAT
TACCAACATTAAAATATACTGGTAAATCAGAAAGTGCAGTGCCAATTGTGT
CAGAAAGTGAATTGCAAACGGTAACAGCAGAGCCGTGGGTGAAAATTTCAG
```
20
```
ATAAAGGGTTACAACTAGAAGGACTTAATTTTAATCGCGAAGGTCAGTTAT
TCTTATTAGACGTGTTTGAAGGGAATATTTTTAAAGTTAATCCCGCAACAA
AAGAGGTTACAACAAAATTTCAGTCTGTTAAAGATAATCCGGCAGCGATTA
AAGTACATAAAGATGGTCGTTTATTTATCTGTTATCTAGGTGATTTTAAGA
CAACTGGAGGCATATTTGCGACAACAGAAAAAGGTGAACAAATAGAAGAAA
```
25
```
TTATTTCTGATTTAAATACAGAATATTGTATTGATGACATGGTTTTTGACA
GTAAAGGCGGATTTTATTTCACTGATTTTAGAGGGTATTCTACACAACCTT
TGGGCGGTGTTTACTATGTAGATCCAGACTTTAAGACGGTTACGCCAATTA
TTCAAAATATTTCTGTGGCGAATGGTATTGCTTTAAGTACGGATGAAAAAG
TGCTATGGGTAACTGAAACTACAACTAATCGACTTCACCGAATCGCATTAG
```
30
```
AGGATGATGGCGTGACTATTGCACCATTTGGAGCGACAATACCATATTATT
TTACAGGTCATGAAGGACCGGATTCTTGTTGTATTGATAGTAATGATAATT
TATATGTGGCTATGTATGGCCAAGGACGTGTATTAGTTTTCAATAAGAGAG
GTTATCCTATAGGTCAAATTTTAATGCCAGGACGTGATGATGGAAAGATGT
TACGTACAACACATCCACAATTTATACCTGGTACAAATCAACTTATAATTT
```
35
```
GTACTAATGATATTGAAAACCATTCTGAAGGTGGATCTATGCTTTATACAG
TTAATGGTTTTGCTAAAGGATATGAGAGTTATCAATTTCAATAAACTCTTG
AAAAAGCGTATAGAATAAGTTGTTATGTATAAATAAAAGAAGTAGAACAAA
GGTTGAATAAAACTTTATGTTCCTCACTCGTATAGCTTAATTTGAAAATTT
GATACAACTCATTGTATTTTCTATTAATGCAACGATTGCCATAAATAAATA
```
40
```
TGCCTGAGACATTTAATGTGTCCCAGGCATTTATAATATGTGTTCTTGTAT
TTTACTGTTAAAGATAAAAGCTTAAGAAAAATTTTCTATAACAAATCGTTTA
ATTGTCGATGCTAAATTCTGCTGATGCTTCAAATTTTACGTTTTTACCTAA
CATCACGCCACCAGTTTCTAAAGCTTGGTTAAAATTAATACCATATTTTTC
GCGGTTAATTGTTCCACTAACGATAAAACCAGTGACTTGTTGTCCATTCAT
```
45
```
TGGATTTTTACTTACACCATTAAATTCAACATCAAATGTCTCTTCATGAGT
TTCACCTTTAATTGTCAAATCTCCAACTACTTGATTTTCGTTAATTTCTTT
AGTTACAAATGTCATTTTATCGTTGTCTTCTGTACCAAAGAAATCGTTTGA
TCTTAAATGGTTGTCTCTGTCCTCATTTTGAGTGTCAATTGAACTTGGAAT
AATAGTAGCTGTTGCTTTTAGTGAAGTTAAATCATTAATATCTCCATCTAA
```
50
```
TTGAACATCGAATTGCTTAAATGTTCCTTTTACTTGGGACACCATAAGATG
TTTAATTTTAAACTGAATATCACTGTGAACTTGATCAAAGTTAAATTTTGT
CATTATTATAACCCCTTCTCTCTTTATTACAATAATTACGTTTCTTGTACT
TGTTATTCTAGATACAGTAAGATGACTTATCAATTTTTTTACTCTCATTAA
ATTTACGGCTTTTTAAACTTTTTTAGTGTGTCTATATTACTATATTCGTTA
```
55
```
AAAAGAAATGAAGTATACATAAGAGAATGATGATATTTGGCAATCAACAGT
AAATATATCTTCCAGAAAAATGTTGTGAAATTTCACATAAATATAAAAAAG
AACATGTGATTGTTGATTTGAGGTTGCTCATAGTGAGGATTATTGAATTAT
AAACTAAAAATATTTCATTTAATACTGAAGTAAAATAACGTATGTCTCAAA
AAATAATTAAACACATACATTTTCATTTGACATTGCAATTAGAAGAAAGTA
```

963

```
ACATAACATTAAATCGTAATTTTTACGATTTAGGAACGATATAGAGATGAT
GGTTGGATGAATATAAGATGGACATATCCGTCCCTGTCTATTTTATTTGAA
GTGCAAACAAAGTGAAAATTCTTCAATACCTCATATAATACATCAATTGAC
CATCATCTATTATATTGTTCTTTTTATGTTTAGTTTCCTTTTATTTCTAAC
TCTTTTTTTTTCACTCGTCTTAGGACGGGTGTTTTTTATAGTAGAGTTAT
GAAATTGTTGAAGGTTCCGTTCGATTATATTCTGAAATTAAATCAATTCAT
AGACTTCTTCGTATCTATGATGACTTGCAACATGGAATTTCAAATACCCCT
TAGTACAATGAACATATTTATATAAAGAGTGAATGGGTTAAACTACGTTTT
TAAAAATGTATAATGAGATGATATTTGAGTTTGTTTTTATATAAAGAGAGG
CAAGTTTGTTTTTTATATAGAAACGGTGTTTATAAAAAATAAATTGTGACG
GAATCGTGTGATTAGTTATTTTAAATTATTGAGTTTAAAAAAGGACAA
```

Sequence 3444
step.1003c08.cons.ok

```
AAATATAGATTGTCGCTTTAACAACAGAATTCAAATCAGAACCTGCTTCTT
TTAATACCACAGTTAAATTTTCTAAAACTTGCTTAGTTTGTTCTTGAACAT
CATCGCTAACAATTGTTCCATCAAGTGTGAGTGGAATTTGACCTGATGTAA
AGACAAAACCGTTTATAACAGTTGCATGCGAATATGGGCCTAGTGCTTCGG
GTACCTTATCTGAGTTACAACAACTTATTCCTAAATATTGTGTTATTTTCC
CTTAAATCTCCTCAAAGATAAATGATTTAAGGGTTTTTTTATGCATAAATC
ATTTGTTCTAACCCGTTTTACTATCTATAAGATTGTTATTGTAATAAAATA
TTGACATTAAATACTAACATAACGAAAAATAATTGAAAGGAATGAAATAAT
GGTACAAAATGCATTTGTTGCACTGGACTTTGAAACTGCTAACAGTAAACG
AACAAGTATTTGCTCTGTAGGTATGGTTAAAGTTATTGATAATCAAATAAC
AGAGTCATTTCATACCCTTGTGAATCCTTTTGACTACTTTACTGAAACAAA
TATTACTGTACATGGCATCCATCCCGAAGATGTGCAAGATGCTCCTGGATT
CAAGCATGTCTATCCATATATGCTAAAATTTATTGATCAACTTCCTGTTGT
TGCCCATAATGCTGCTTTTGATATGAATGTGTTACATCAAAGTTTAAAAAG
TCATAATATAGATACACCTTCTTTGACATACTTTTGTTCATATCAATTAGC
TAAACGAACAATTAATGCATATCGATATGGGCTAAAACATTTAATGAATCA
TTACCATTTAGATTTTCACGGTCATCATGATGCGTTAAATGATGCTAAAGC
TTGTGCAATGATTACCTACCGCTTATTGAAACATTATGATGATCTTCAAAG
CATGTTGAGTATATATGGAAAAAATCTCAAAGATAAAGGCTGATTGTCCAC
TTAACTTGTGTTCATCAGCCATTTAAATAAAATCATATATTGTTAAATTTT
TAAAATCAGATTGCTCTAAACTCCCGACCGTTACACCTATTAATCGAATCG
GAATCTCTGGATCTTTTAAGTCGTTATATAAAGTATAGGCAATATTATAGA
TATCAGTTTCAGTACGAATAGGATCTCTCAAACTTTTTTGCTTTGAGATTG
TTTCATATTGATAGGTCTTTATTTTTACAGTGACTGTTTTACCTGATTTTT
GAATTTTATTTAAACGTTCTGCTGTTTTTCCTGAAAGCTCTCTTATTTTTC
TTAAAATAACATCATCATCGTTAACATCTGTAGAAAATGTCCTTTCAGTCC
CAACAGATTTACGAACTCTACTCGCCTTTACTTCATTATGGTCTATACCTC
GAGCTTTATTATACAAACCTCGTCCTCTTTTACCGAATAAACGTATTAATT
CAAATTCATCTTTATTATATAAATCTTGACCTGTATATATATGATGTTGAT
GCATTTTTTTCTTCGAAGCTTTGCCTACCCCTGGAAAATCTCCAATATCTA
ATTGCATTAATATTTCATGTACATTATTGTAATCAATTACTGTCAAACCAT
TCGGCTTGTTCATACCACTCGCTAACTTTGCTAAAAACTTATTATAAGACA
CGCCAGCTGACGCAGTTAAACGTGTTACTTCGTATATATCTCTGCGAATAT
AATTTGCAATGGTTGATGCTGGTAAATCCGGTCTCACTAAATGTGTAATAT
CTAAATAAGCTTCATCTAAAGACATGGGTTCTACTAATTCTGTATAACTTC
TGAATATTTTCATGATTTGACCAGATACCTCTCTATAAGTATCAAAACGGC
TTGTTACATAATATCCATTGGGGCATAGCTTATGTGCTTGAGTCATAGGCA
TAGCAGAGTGAACACCATAAGCTCTTGCTTCGTATGATGCCGTAGAAACTA
CGCCTCGATGACTCGCTTTACCGCCAACGATGACAGGTTTTCCTTTTAGTT
TAGGATTATCTCTCATTTCAACTTGAGCAAAAAAATAATCCATATCTATAT
GAATAATTCTTCTTTCAGTCATCGACTCACCTCACTATGTCTTCCAAATAA
TGGCTTTATATTATTATACGCCATCCTTTCTATGTTTACATGTATTGTGCA
CATGGTAATATTTGTATTGTTTATTTAATTAATGTAATAAAAAAAATAGGA
CTACAATTTTTAGAATAAATTCATAGTCCTATCCTTTTCTACATAAGTAGT
```
```

TATGTGAATGTAGTTGCATATTTTAATGAGATTTTTTATTACAATTGAAAC
GAGTTGAAAAATAAGAAATGAAATTTTTAAATCAATTATCAACTTTGATAT
ATACGCTTGTCACTATAATTAGTGATTATTCATAATAGGTTTGAGTAGGCT
TTTGATGATAAAGATACAAACCTTCTCCATAGGTTTGTATGTATTTGAATT
5   TGCTTTGTAATGCCAATGCTACTAAAAACACATCTATCATACAATAACTTG
TATGTATACTAAACATAAAGATTAAAGATGAGAAGGAATATATTTGAATAA
GTGCTAGTAAGGTCGCACTTATGATAACAAGTGGTGCTATCATGATAATTG
TAAATTGCCACCGATGAAAACAAGTCTGAGAGAACTGTACAATAATTTTGT
TTTTATCATATTTTAAAGATGGTTTTTCACCTTTAGAAAATATAATGAACA
10   AAATACGATGAATAAACTCGTGCAAAATAACTAGTACTGCGAAACCTACAA
ATCCAAAAATAAGATTCATGACAAGATTTTGTTCAATGATATGCGTTGTTT
GATAAGCCCATTTATAAGTAAATAGTATTGTAAATAGCGCCAATAC

15   Sequence 3445
step.1003c09.cons.ok
AATAAATTTTTTGAAACGGCTTTGGAATTGATATTCAAATATATCAAAATC
TAAACCTACCAATAAATAAAAAATTGCCATACAAATTGCCATAACGACCAG
TATACTTATTTTTTTATTAGCGCTCATTTGCATAATTTTTCCTACCTTTCA
20   TTAGCAAGATTAGGAAAATAATTGTGCCAAATACACCTATCGTCAATCCAA
TATTAATTTCGTAAGGGTAAACAATTATTCTACCAATAATATCAGCAATTA
AAACAAAAATTGCACCGAGCATAAGCGTGTGTGGTAAAGCATTTTTCAAAT
GATCGCCTCTATAGATAGATATGATATTCGGTACGATTAAACCTAAAAATG
GTAATGTTCCAACAGTAACAACAACTAATGCTGTTAGCGTTGCTGTTATGA
25   ATAATGCTATTTTAATGATTTTTTTCATAACTTACACCTAAATTATGACTAA
AGTCTTTTCCCATACCTGCAATAGTAAAATGATTTGCAAATACAAATGCCA
AAATAAGTAATGGTATTGTGAGATACAACACCTCAAAACGACCACTCGTTA
TAACTGCAAAGTTACCAGTTAACCAGTTTCCAATGCTTTGTAAAGCATTGG
TTCTCAACGCTACAAATGTAGTAAAACTGGATAAAATACCACCAATCATAA
30   TGCCTAAAAGTGGAACAAAGATTACATCTTTTACACGGATAAGATTAATTA
ATTGGACAAATAAAAACGTTCCAACAATACTTAGAACAACAGCAAATAATA
ATTTGATTAAAATGGGACCATTAGGAAAGAACAACAATGACATTAAAAATAC
CTAATTTTGCCCACTCCATAGTACCAGCAGTCGTTGGGCTTACAAATTTAT
TTTGCATCATCTGTTGCATAATTAATCCTGATAAAGCTAGTGAACTACCCG
35   AAAGTAGAATACTAACTGTTCTGGGAATTCGACTCGAAAACAAAATATTTA
TTTGTTCATCACTTAAATGGAAAATATCTATTAGAGAGAGCTGACTCACCC
CTATAAATAAAGAGACAATAGTTAAAATCACTAATAAAATAAATAAGGTAT
AACCTTTAAAAATAAATTTCATTAGTTAATCTCCTTGCTAAATATTAAAGA
TGTATAAGGTCAATATACGAAAGTAATGATAATGATAATGATTATCATTAA
40   CATCAATATTATAGACTCTATTTTCTTTTTTCAATAGTTTTTTCATATTTT
TTGGGAATTTTATAAATGTTTTATCCAGCAGTAATAATACATCAAAATATT
TTATAATACCAACTTTCATAAAAAAAGAGATTCCTACTAATTTTAATAGAA
ATCTCTTTAAAAGTACTATATTAATTATCAAAAACAAAGTCTTCATCACGT
AGAGGCTCAACATTTAAAGCCAATGTATAGCCATCACCTTTAACAGAGAAG
45   AAATCGTGGTTCTTTGTAGATGTATCTAGTGCATTTTCAATAATAGGGTTA
AACTCTCTCTCTTCGAAATATGGTTCAAAACCTAAATTCGATAATGCTTTA
TTTCCATTATATCGAACATAATTTAATACATCTTCAGCAAGACCAATATCA
TCATATAATAAATGTGTATATGAAACTTCATTATCATAAAGTTCATTTAAT
AATTTGTACATTTCTTGATCAGCTTTTTGTTTCTCACTTTCAGATAACTCA
50   TTACGTAGACTTTGTGCATCTAAACCTGTGAACACTCCATGTATAGATTCA
TCTAAAAGTATCTTACGTATAATTTCACCTGACGTAGTCATTTTTCCTTGT
CCTGCGAGATATAATGGATAATAGAAGCCAGAATAGAATAGAAATGTTTCT
AAGAATACACTAGAAACACGAGCAATATATTGATCGTAAATCGATGCTTCT
TTACCCCAAAGTTTGTGGTAATTTTCTACAATTTTATCTGATTTATATTTT
55   AAATGTGGCTCTTCAATAACCCAAGTATCCAATAAATAGTTGGTTTCACTA
GATGGTAATAATGTAGTGAAGATATGAGAATAACTTTTCGCATGGATTTGT
TCCATCATAGCCATAAATGAATAAACAGCTTTCTTTCTTAAATCAGTAGTA
TGAAGCATGATTAATGGCATACCATCATCAGCTTGATGTGTATCTAAACCT
GTTAAACCTGCAAGCGCTTTTTTTAAAAGTATTTTTCTCAGAATCTGTTAAT

GTTTTCCAACTTGCTATATCTTTTGATACTTTAAATTCTGTTTCCACCCAC
ATTTGTGAGATATTTTGACGCCAGAACATATTAGTCATATCTTCTTGGGTG
TTCCAATTTACCGCCTTCATTAAAATCCTCCTGTCATATTTATGTATAAAA
TAAAGAGTTTTGAACCAATGTAAATCTATAATTTAAAGAAATACATCAATT
5 CATCACTCTACTCTAAAATCTTTACAACTTAATCAATTAAATCGCACAACT
TGTACATTCTTCCACACTTAATAATTTATTACGTGTATAGTATAAAGATTT
AAGGCCTTTATGGTGTGCATATACATATAATCTTGAAAGTTCACGTGTTGA
AATTTCTGAATTAACATAAAGTATCGTTGAAATACCTTGGTCTACGTGAGT
TTGAATTGTAGCTACTAAGTCAATAAGTTTCATTTGGTCAGTATTAAATGC
10 TGATTTGTAGTACCACATTGTTTCTGGTGATAAAAATGGCATTGGATAAAA
TGTTTCCGCATTTCCATACGTACGACGTTCAATTTGGTCAACAATAGGCAT
AACAGAACTTGTCGCATTTTGAACATAAGAAATACTTTGAGTCGGAGCAAT
AGCTAATCTATAAGCGTGATAAAGTCCATATTGTTCTACTTTATTTTGCAA
TTCTTTCCAATCATTTGAAGTAGGTATATCGATACCATCAAAAAGTTGGCG
15 AACCTTTTCAAATTTTGGTTCAAATTCTTGAGATGTATAGAATTCAAAATA
TTTACCATTTGCATAGTCTGATTGCTCAAAGTCTTGATACTTTTCTCCACG
CTCTTTTGCAATTTCCATTGAACGTTCGATGGAATAATAGTTCATAATCAT
AAAGAATATATTAGCAAAGTCTTTTGCTTCTTCAGATTCATAGCCAATTTT
ATTTTTAGCTAGATAACCATGTAAGTTCATTACTCCTAGTCCAACAGAATG
20 TAGTTCACTATTCGCTTTTTTTACACCTGGTGCATTTTGAATATTTGCTTC
ATCACTTACAACTGTAAGAGCATCCATACCTGTGAACACAGAATCTCTGAA
TTTACCTGATTCCATAACATTCACTATATTCAATGAGCCTAAGTTACATGA
AATATCTCTTTTAATTTCATCTTCAATTCCATAGTCGTTAATTACTGATGT
CTCTTGTAAT
25

Sequence 3446
step.1003c10.cons.ok
TATATTTCGTGTGATTTTCTTTAACTACTGCACCTGCCACAATTCGATGCG
30 GTTTTTTCTTTAATTCTTTTAGTAAAGTAATTCCACGTTGTGCGCGTTTAG
CTTCTTGAAGAACATTAAAAATCAATACGCTTCATAGCACCACGTTGTGTAA
CCATTATTATGGAATCTGAGTCGTTCACATCTTCTGTCATAACAACATAGT
CTTCATCTTTAAGATTAATTGATTTAACACCAGCTGCTCTTAAGCCTGTAT
CCGATAATTCATTAGTTGAATAAGTTAATGACATGCCTTTATGGGTTAGAA
35 CAGTAATTAACTGATCAGACTCTAATCGAACGACATTAATAAGTTCATCTT
TGTCTTTAACCTTCATATTTATGAGTGGTTTATTAAACCGAGTAGTTTTAA
ATTGTGAAGCACTACTTTTCTTAATCATGCCGTTTTTTGTAGCCATAATAT
AAAAGGCTTCATTTTTAAAATCTTTTTCGTTGTATACATTTACCACTTCTT
CATCTTCATCTATTGGCACAATTTGTGATATGTGTTGACCAAGCTCTTTCC
40 AACGGATATCGGCTAATTTATGAACAGGTATAAACAAATATCTACCTTTAT
TTGTAAATACAAGAACAGTATCTTGAGTATTCACGCTTTCATGTTTTAATA
AACGGTCGCCGTCCTTCAAACCGATTTCAGTCACACCACTTGCGTTAAAAC
TACGTGTAGATGTACGTTTTATATAGCCATGTTGCGTCAAACTTAAAATCA
CTTCTTCACTAGGCACCATAACTTCTTTATCAATTTTGATTTCGGAAATTT
45 CAGCTTCGATTGTAGATAGTCGATCCACTTTAAATTTCTTTTTAATTTCAT
TTAGTTCATCTTTAATTACTGCTAAAAGTGCCTCATGATTATCTAAGATAT
TTCTTAATTCTTTTATTAAAGCTTCTAACTCTTCATGTTCTTTTTTCAAAG
CTTCAATGTCAGTATTTGTTAATCTATACAGCTGTAACATGACAATAGCTT
CTGCTTGAGCTTCAGTAAAGTCATACTCTGCAACTAAATTATCTTTAGCAT
50 CTTTTTTATTTTTAGAATTACGTATCAATGCAATAACTTCATCAAGTATAG
ATAAAGCTTTCATTAATCCTTCCACAATATGCATACGTTTTTCAGCTTGCT
CTAAGTCATAACGCGTTCTATTTGTAACCACTTCAATTTGATGATTTAAAT
AACTTTCTATAATTTCACGTAATCCCATCAACTTAGGGCGACCTTCACTAA
TAGCAACCATATTAAAATTATATGAAATTTGTAAATCCGAATTCTTATATA
55 AATAGTTTTTGATTGATTCGCTATTAGCATCTTTTTTTTAATTCGATTGCAA
TTCGTAATCCAGTTCTATCAGTCTCATCTCGAACTTCTACAATACCATCAA
CCTTTTTATCGGCACGTAATTCGTCAATTCTTTTAACTAAACTACTTTTAT
TCACTTCATACGGAATTTCAGTCACAATTAATTGTTTACGTCCACTTCTTA
AAGGCTCTTCATCTACTCGTGAACGCACGACAACCTTTCCTTTACCGGTCT

```
     CATACGCTTTTTTTATACCTTCTATTCCTTGAATGATACCACCTGTAGGAA
     AGTCAGGCCCTTTGATATATTTCATCAGTTGATTAATTGTAATATCAGGTT
     GATCGATATACTTCAATGTGCCTTGTATTACTTCGGCGAGGTTATGCGGCG
     GGATATCAGTAGCATATCCTGAAGAAATCCCCGTAGATCCATTAATTAATA
5    AATTAGGAAATCTCGCTGGTAATACCATTGGTTCCAAAGTTGTGTCATCAT
     AGTTTGGAATAAATGATACTGTTTCCTTATTAATATCCCTTAATAGTTCTT
     CTGATAATTGACTAAGTTTAGCTTCTGTGTAACGCATAGCAGCTGGAGGAT
     CGTTATCGATACTACCATTATTACCATGCATTTCAATTAGAACATGACGTA
     ACTTCCAATCTTGACTTAAGCGCACCATAGCATCATATACTGAAGAGTCTC
10   CATGAGGATGATATTGACCTATTACATCACCGACAGTTTTCGCACTTTTAC
     GGAAATTTTTATCATACGTATTCCCACTTGAATACATTGCGAATAGTATAC
     GTCGTTGTACTGGTTTTAAACCATCACGTACATCAGGTAATGCACGTTCTT
     GAATGATATATTTACTATAACGTCCAAATCGGTCACCAATAACATCTTCAA
     GTGACAAATCTTGAATTATCTCACTCAATTCGTTTCCTCCTCAATATATTT
15   TTCATTCTCTAGTATTTGGACTTCTTTATTATCCAAAATGCTTTGATCTTC
     TTGCATACCAAATTCAACGTGTTTTTCAATCCACTCTCTTCGTGGGGCAAC
     CTTATCCCCCATCAAAGTAGTGACACGTTTTGATGAACGAACTTCATCTTC
     AACTTGAACTCTAATTAATGTCCGAGTTTCTGGATTCATGGTAGTTTCCCA
     TAATTGTTCTGGATTCATTTCACCAAGACCTTTATAACGCTGTAATATGAA
20   ACCTTTTCCTAATTGCTTTTGTAAATTTTCTAATTCTTCATCAGTCCAAGC
     GTACTCAACTTTTTTATTCTTACCTTTGCCTTTTTCTAATTTGTATAAAGG
     CGGTAACGCAATAAAGACACGTCCAGCTTGAACAAGTGGTTTCATATATTT
     AAAGAAAAATGTAAGCAATAATACTTGAATATGTGCACCATCCGTATCAGC
     ATCTGTCATGATAATAATTCTGTTGTAATTACTATCCTCAATTTTAAAGTC
25   AGTACCAACACCAGCACCAATAGTATGAATAATCGTATTAATTTCTTCATT
     TTTAAAAATATCCTCTAAACGTGCCTTTTCTGTATTAATAACCTTTCCACG
     AAGAGGTAAAATAGCTTGGAATTTACGGTCGCGTCCCAATTTTGCAGAACC
     TCCCGCTGAATCACCCTCAACTAGATATAACTCGTTTTTATCAGTATTTTT
     ACTTTGCGCAGGAGTTAACTTACCTGATAACAATGTATCTTTACGTTTATT
30   TTTCTTTCCGGAGCGTGCATCTTCTCTAGCTTTACGAGCAGCCTCGCGTGC
     TTGTTGAGCTTTAATTGCTTTTTTAACTAATGATTTAGATAATTGGCCCTT
     TTCTTCTAAGTAATATGGTAATTTTTCTGAAACAACAGAGTCTACAGCACT
     CCTTGCTTCTGAAGTGCCAAGTTTTGATTTCGTTTGCCCTTCAAATTGAAG
     AAGTTCTTCTGGTATACGTACTGAAATTATCGCTGTTAAACCTTCGCGTAT
35   ATCATTACCGTCT


     Sequence 3447
     step.1003c11.cons.ok
40   CAAATTTACATCGACAAACACTCTTCGATGAAGAAGACGCACATACTTTAA
     TATCCAAAGTTGAAGCAATCAAGCATAAGGTTAATCAAATTAATATAAATG
     TCAATCTAACAACTTATGATTTAGAAGTTACTTCATCAAATATCGAAAATT
     TGATAAAAAATGTCGAACCAGACATCATCATTGATGGCATGGATAACTTCA
     AAATACGATACCTGATTAATGAGGTTTGTCACAAGTATCAAATCCCATGGG
45   TTTATGGTGCAGCTGTTGGTAGTAAAGGATCAGTATATGGAATAGATCACC
     AAGGACCATGTCTAAAATGTTTATTGCAAACAATTCCTGACACAGGGGAAA
     GTTGCGCTATTAATGGCGTAATTCCCCCTGTTATATCAATGATTGCAAGCT
     ATGAAGTAGCAGAGGCCGTACGTTATCTTTCAGGAAAAGGATTTTCAAAGC
     AATTAATCACTATTGATGCATTTAATATCAATTATAAGTCAATGAATGTAG
50   ATGCACTCAAAAATAAAGATTGCCCAGTGTGTGAAAAACATGAATATACGT
     TACTAGAAAGCCAACAAGAACGTACTATTGAGGACTTGTGTGGGAATGCTT
     ATTTATTTAGATTCCCACCTAAAGCTTTTAAACACGCTGCCCATTTCCCTG
     GGAATATGGTGAAATCTACTTCCTTTGCCAAATTAATTCAATATCAAACTT
     ATGAATTCACCTTGTTTAAAGATGGTCGTATGAATGCATATGGTATACACA
55   ATGATGAAGAAGCACATCACCTATACAATACGTTGTTAAAATCCATACGCT
     AATTAAAGTAAGATTGTTACCCTTTCAACCTGTTTCTTGTTAAAATAACTT
     ATACTTAACATCAAGTTTACGACAACTGGATTAATTTAAAAGGAGAAACTA
     ACGTTGTATGGAACAAATTATCACTGATTTTATTAGTAAGTGGGGTTATAC
     AGCGATATTCATTTTAATCTTATTAGAGAACGTATTACCTGTCGTTCCATC
```

```
     TGAGATTATTTTAACTTTTGCAGGCTTATTATCTGTGAAATCACACTTATC
     TATTTGGACATTATTAATCATAGCAACAATTGCTTCATTCATTGGTTTACT
     CATTTTGTATTATATTTGTAGACTTATCTCAGAAGAGAAATTATATCGTTT
     CGTTGATCGACATGGTAAGTGGATGAAGTTAAAAAGTAAAGATTTGAAACG
  5  GGCAAATGATTGGTTTAAAAAGTATGGTGCGTGGGCTGTATTTTTATGTCG
     TTTTGTCCCAGTACTTCGAGTATTAATTACAATACCTGCTGGCATTAATCG
     AATGAACGTTATACAGTTTACAACTTTATCTTTAATAGGTACTACAATTTG
     GAATTTTGCTTTAATACTGCTCGGTCGTTTGCTCAGTGACAGTTTTGACGC
     TTTGATGAATGGTATTCATACATATTCACGTATCATGTATGTCATTATTAT
 10  TATTGCAGTCATATATTTTGTTATACGTTATTTAATGAAACGTCGTCGGAG
     TGTTAAATAAATGTTATGCTAAGTTTTCCGTTTCTTAACGGTTAACTTAGC
     TTTTTTATTAATTTTAATTTTATGATACGCAATTATTGAAAATTAATTGAT
     TTTTCAAGATAATCTATGTAAGCGCATTATTATAAATAATCGTATCGTTAG
     GTCGGAAAATATAATTAGAGGTGACTATGCATGACCCATATTACAGAAAGT
 15  GAAATGAAACAAAAATATCTAGATTTACTCTCACAAAAATTTGACAGTGCA
     GAAAAACTTGCTACTGAAATTATTAACTTAGAGTCAATCTTAGAATTACCT
     AAAGGGACTGAACATTTTGTTAGTGACCTTCATGGTGAATACGAATCTTTC
     CAACATGTTTTAAGAAACGGATCTGGAAATGTGCGTGCTAAAATTAATGAT
     ATCTTCAAAGATAAATTATCCCAGCAAGAAATCAACGACTTAGCAGCATTA
 20  GTATACTATCCGGAAGAAAAACTAAAATTAGTTAAAAATAATTTCGATTCA
     ATCGGAACATTAAATATTTGGTATATTACAACCATTCAACGATTAATTGAT
     TTAATTACATATTGCTCATCAAAATATACACGTTCAAAATTACGCAAAGCA
     TTACCTGAACAATACGTTTATATTATTGAAGAGCTACTTTACAAGAGCAAT
     GAATTTCATAATAAAAAGCCTTATTATGAAACATTAGTTAACCAAATTATT
 25  GAATTAGAACAATCAGATGATTTAATCATTGGCCTTTCCTATACTGTACAA
     CGTCTAGTCGTAGACCATCTTCATGTCGTGGGCGATATCTATGACCGTGGT
     CCTAAACCTGATAAGATTATGGATACATTAATAAATTATCATTCTGTAGAT
     ATCCAATGGGGAAATCATGATGTATTATGGATTGGCGCCTATGCTGGTTCA
     AAAGTATGTCTTGCTAACCTTCTACGTATCTGTGCACGTTATGATAATTTA
 30  GATATTATTGAAGATGCATATGGCATCAATCTACGCCCTTTACTTACGCTT
     GCTGAAAAGTATTACGATGCTGAAAACCCAGCGTTTAAACCTAAGAAACGA
     CCAGATAAAGACGTCAGTCTTACAAAACGCGAGAAAGTCAAATCACAAAAA
     TTCATCAAGCAATTGCGATGATTCAATTTAAACTTGAAATGCCTATTATTA
     AGCGTCGTCCTTCTTTTGAGATGGAAGAACGTTTAGTTTTAGAAAAAATAG
 35  ATTATGATAACAATGAAATCACTATATATAATAAAACATATCCGCTTAAAG
     ATACTTGTTTCCA


     Sequence 3448
 40  step.1003c12.cons.ok
     TGAGTTCTAATTGGCAATGATTTAAAACAGATTGAACAATGTGAGACACAT
     ACTGTTCTTGTATTTCTAAGCTTTCTGCATGATTTGTAAAGGCCATTTCTG
     GTTCAATCATCCAGAATTCAATTAAATGACGGCGTGTTTTAGATTTTTCTG
     CACGAAAAGTTGGGCCAAATGAAAAAACACGTCCGTGAGCCATTGCGGCTG
 45  CTTCCATATACAACTGCCCACTTTGTGATAAGAATGCATCTTCATCGAAAT
     ATTTTGTATGGAATAACTCACTTGTTCCCTCTGGTGCACTTGCTGTTAAAA
     TAGGTGGATCAATTTTAGTGAAGCCATTTTCATTGAAAAACTCATATGTTG
     CACGGATAATTTCATTTCTTATTTTCATGACAGCATGTTGTTTTTTTGAAC
     GTAACCATAAGTGACGATGATCCATTAAAAATTCTGTTCCATGATTCTTTG
 50  GTGTAATAGGATAATCGTGTGCTTCATGTACAATTTCGATTGATTTAACTT
     GCATTTCGTAGCCTAAATCAGAACGATTATCTTCTGTAATCGTTCCTGTGA
     TGTATAAAGATGATTCTTGAGTTATATCTTTTGCTAGTTGAAATGTTTCTT
     CATCTACTTCAGATTTTACTACTACTCCTTGCATAAATCCTGTACCATCGC
     GTAATTGTAAAAACGCTATTTTACCACTTGAACGTTATTAGTTAACCAAG
 55  CACCAATTGTTACTTCTTGGTTAAGATGTTTTTTCGCTTGTTTAATCGTAG
     TTTTCATAACCAAACTCCTATATTTTTCTTTTTAAATACTTATTTATTTTA
     ACAAAATATGTCTTTAACTTCTACAAGTTGAGTAACTGGAATCATTCGATA
     TTATTTATTATATATTTCTAAAACGTTAGCTTTTACTCTCATTATCGATAG
     ACCTCTGTATTTGATTCAGTAATTTATTAAACTGTTTAATATTACCTTTTC
```

GTTGTTTAAAATGTTCTAGTGATTGAGCAAAAAAGCTTTTATATGTGCTAT
TTACTAGTCGATCATCAAATGACACAATAAGTCCTTTGTCGTCTTCGTGTC
GAATAAGACGTCCAAGACCTTGTCTAAACCTTGTTACAGCATCAGGTAATA
CATAATCTTTAAAAGTTGATGTGAATTCAGAATCCATAAGCCAGTATTTAA
5   TATTATGTTTATTCATAAATGGTAATTTAGCTATCATTACACATTTAAGGC
CATTGGCTTGAAAATCAAAACCTTCAAAAAATGTTGATGTTCCAAGAAGTA
TAGATTTATCAAAGTTATTAAATTGTTGAACAATTTTATAATTTTGGTTTT
GTTGTTGGGTTAAAATGACATAATCTTCTAACTCAGGTAATTCATTTAACA
AATCTTGAACCATATGCATCATTTTATAGCTTGTAAATAAGACAAGACATT
10 TCGACTGTGTCACTGTGATGTATTCAACAATATAGTCTACAATTGAAGCTA
CATAATCGTCTAAATTTTTATAATTGTACGTTTCAACATCATTGGGCACAA
AAACATTTGTATGGTTGGATGATGTTAAAGGTGTTGAAATTTCAAATGTAT
TAAAGTCTATATCTTCATTAAACCAATTTTGAAATGCTTTAAATGAGTGAT
TGAAAGTTAATGTGCCAGATATAAATGTAAGAGACTTAAATTTTTCGAGTA
15 CTTGTTTCGTCAAGATATCTTTAACATCATAATCCTTGACTAAGAGTCGGA
TTGTTGATTTTTGTGCTAAATTTTTAATTGAAATAAAGCTTGTATGATGAT
CTTTTATACTTTGTTCGATAAGTTTAAATTTGTCATGTAAGTATAATAATT
GTTTCCGTACAGACTTGATTGTTTTGTGATTCATTCCGTTAAAAATTTCTA
TCGTTTTATTTAATTTATCTATGATTGCACGTAAATCTTTTAAAATCTCAC
20 CCGTTTCAAAGTCATAAACGTAATGATACTTATGAATGTCATCATCATAAA
CGTCTGATGTTTGAATAATATTATAAATTGTAGTGAATAGTTGCTCATTTA
AATCATGTAACTCATTGATATTTATTTTCAGTCCAAAAACGTCTATAGGTG
CAATATCTAGTTTCTCTAAAATACGTTGTTGCTCAAGTTTGTCTACTGCTT
TAAGTAGTTTTTCATTTTCATTTTTGCCAATAAGTCCTAATTGATATTTAA
25 CATCTGAATAATTTAAATCATTAGTAACTTGATTTAGCGCATAGTCAGGCA
ATCTATGTGCTTCATCGATGATGCAATCATCAAATAGTTGATATATAGTGT
TTTCACTGTCTGAATGAATTAAGTGCGCATGATTAGTGATACCAATTTGTA
TGTTTTGAGCATTTCTTTTTATATAATTATAATAATGGATATCATGACGAA
CTGGAACGTATGTTTCAATTTTTTGGTCAACATACATTTTTTGTCCACCTT
30 TAAGATTTAATTCCTGTATGTCCCCAGTATTTGTTTCAGTTATCCAAATAA
GTAACTGCATTTTAAGAATACTTACTTCATAATTATTTGTATCGTCTTTAA
GAATTTGGCTGATAAGACCAAGAGATATATAATCATTTTTACTTTTGATTA
ATGACGCGTTAATTTTAAAATCTAAAACATCATTGAGTAATGGTATGTCTT
TCTCTAATAGCTGACTTTGTAATAATTTTGTATTTGTTGAAATCATTACAT
35 GACGACCGGTCTCAATATTATACATTGTTGCTGCAAGCAGGTAAGCAAGAG
ACTTTCCACTACCCAAAGGAGCTTCAATCATTGCCTTATCACTATGCATCA
ACTGATCAAGTATAATTTCGGCTAGGTATAACTGTTGTGGCCGATATGTAA
GATTTAATGATTGAGTGACATTTTTATATAAATCTTTTAAGGTACCATCAA
AATTGACAGCCGGTTTTTTTAAATCAATTTGTTTACGGTAAATAATTTGTT
40 CAAATTGTTCAAATTGGTTGTTAGGCGGTTTAGTTTGATAATTTCTAACCA
TTTCAAACAAAATATGATAAAGATCATACTTGAGATTTTTACTTAAATAGT
ACAGTTGTTTTTGTGTATCTAAATGCAATTGCTCGAACTTCTCAAAAGCTT
TAATCATCAATTTAGCAGTTGTTGTTGCATCTTCATCTGCTCTATGTGCAT
TATTTAATGGTATATGATGAGATTCAGCCAATGCACTTAACTGGTAGCTTT
45 TGTCTGTAGGGAATGCGATTTTAAACAATTCTAAAGTATCCATTACTCTTT
TAGGTTTAAATTGAATATTAC

Sequence 3449
50 step.1003d01.cons.ok
AGCTTCAGCAATCAAAAATGGTTGGAAAATGAATGACTTTGTAGGACCTGT
TGATCATGGTGCATATAACCGTATTGAACAGATTAACGTTGAAGTCACTGA
GGTTTAAAATACGATATATATAAATACTAGTACATTTTACCCCTCATCTAT
AGTCTCATTTAATTCCTACAAACATTGTAATTATAATTAAATCAATATCTA
55 TATAGATGTGGAAAGCCTGGAACAAAAAGTTCCAAAATAAGTAAATAAAGA
CAATTTCTATTGAAATAATATAGGAATTGTCTTTTTTATAATTTTTTTGAT
GATTTTCAGCTCGTTGAGCTATTACTTTTCTTATATTAAGCCACGAGGATG
TCTGAATTCATCGAATTCAGTATCGTGAATTGTTTCTACAATATCATGTAC
ATATCGTGAAATATCATTCGCGGGGATAAGTACTGAAGTTTCTATTGGTAG

```
      AGTATGTTAAGTCATATTATAATCTTTATACATAAGACACCTCGTTAATTT
      AGTTTAGTGGTATTTATTAAATTATACGAAAGCACCTTATTTTTCAAAGTA
      TTTTAATATAAAATTACATATAAACACAAAGTATTTTGGCGAGACTCTTGA
      GGGAACAGGACAAGCTGAAGACTACAGGCTGAAGCTGTCACCTAAGAAAGC
  5   GAGCCAACAATACGAATTATTTTAAATAAAGATGCCAATAAATGAATTTAT
      GAAAATCCATTTACTGGCTATTTTGCTAGGAATTATGTCTCAGCGTCTCTT
      CTCTTTCCTTTCAATCCATAATAAATATTGTATACATATAAAGTCTTAGAA
      AGTAATCAGTTTATCGTCGTATGTAAATTTGATACTACAGACCCATGATTT
      CTTATGATTTATCTTTTATTATTTAAAGTGTAGAGAGGCAAGGGAGGTATA
 10   AACAACATGATTATCTACAGAAGAAATATAGAAAATGGAACACCCATTTAT
      GAAATCATAACTAAAACTTTCAAGACAATTACTATAAAGTGTGATGAAACT
      TTTAATAAGTATGAAATCTATCAATTGCTCTCTCTACTAGAGAATGACGTT
      GACAACATGCCGACAAGTTACTCATATCGTTAACGTTATTCTAGACAACTA
      GACATTGCATCAAACCTAATTGAGATTTAGCACTTTTGTCTTTCCAGTCCC
 15   AAATAAAATGAGTTTTAAAGCATTTAGTCCCTTAATACGTTTAGCTTTAAA
      GACATTCTATCATCCCTAAATTAACACTCACCTTAACAAGTGAGAAACGCT
      AATGTTAGCCAAAACTTGTCCCCTTTTGTAGCTCGGAGAGAACCTTTGAGC
      ATTGCTAACATTGGCGTTTTTAATTATTTAATTTAAAAAAAGATTATTATT
      TTTATCATTCTCAACTTCTAGTAATTGTTAGTTTGGTATCATCATATGTTT
 20   TAAAAACTACAATTTCATCATTTTCTTTTATTAAAAAATTGTACTTCTTCAT
      CTATATTTATAGGTTTTAAGTACTTAATTTTGTAATTCATCCAATTCACAT
      TAAGTTCAGAAAATACTTTTTCACATATGAGTTGTCCTGGTACAATGGAAC
      GATGGATCGGATTCTCATCATTAACTATAGTTAAATATTCATTCACCATTT
      CTATTGTAAATTTCATAATTTATCTTACCTCGATAAATATTTGTTTTATAT
 25   ACATACAATGTTTTTTATCTTTATTGATTTCGAGTCCATATGTATATTTAA
      TGATATTTTTAATTAATTTTTGTGAAACCTTATGCAATGTCGCTTCATACG
      TACAATCGACTTGAAGTTGATGTTGTGTCTTAATGTGTGTCTCTTTTAATA
      TGATAGGCTTCTTTGAAAAAGATTGAAATAACTCAAATTGTGGCCATAACT
      TGGCACACATAAGAGTAGGAACATAATGGTCATACTTTATATTAAGCAATT
 30   CGCAATAATTTTGAACTTCTTCATTTGTAAATTGAACTTCTCTAATTTCCA
      TACCATCTTTCAAATAAAGCTGCATTTCCTATCCCTCCACCTATTCCCATA
      GTTGCTATAGTTCTAAATTGATGTTTCATATAAAATAAACGTGTGACTAAC
      GCTGCTCCACTTGCACCATATGGATGCCCTGTAGCAATTGCTCCTCCCCAA
      CAATTCAACTTATTCAAAGGAATGTTAAGCTGTTGTTGGCTCGCAATAACT
 35   TGAGAGCTAAATGCTTCATTTAATTCTACTGCATTTATATCATTTATAGTT
      AATCGTTCTTGAGCTAATAATTGATTTACTGCTGGCACTGGACCTACTCCT
      AAATACTGTGGTTGAACACCTACAGTTGCACTATTCACAAACTTAATCCCT
      TCTGTGAATCCTAATTGACGTGCCCGATTTTCTTCCATAACAATCAGTAAT
      ACTGCACCATCATTTTTCTTACAACTATTTCCTACTGTGACTGTTCCTTCA
 40   TTTAAAAGTGGTTTAAGTCTGCCAAGTGTTCTGAGAGTAAGTTGAGGTTTA
      ATACTTTCATCTTGATTAAAATATTCACCTTTCACTTTGAACGGTAAAATT
      TCTTGGGAAATATTACCGTTATTCATATTTTTTGATGCCAACTGATGACTG
      CGATACGCAAAGTCATCTTGTTCATTTCTACTGATATGATATTTCTTCGCT
      ACATTTTCGGCTGCTTCAATCATTGAAGGGTCTTCTCCTTCTCTTGCAAAA
 45   GGCGCCCGTTCAAAAAATTGTGGAAACTCAGATTCATAAACTGACTGCGGA
      CGTTTGATTTTCCAAGGTGCTCTACTGGTACTCTCAACACCACCTGCAATA
      TATATTGTTCCAGCACCACTTTGTACCATCCTACAGGCTTGTATAACGGCT
      TCAAGACCTGAGCCACATTGACGATCAATTGTTATACCAGGTATTTTAAAA
      TCTAATCCCGCTTCAAGTAATGATTTTCTAGCTAAATTCCCCCCATTACCT
 50   ACCGTATTACCTA


      Sequence 3450
      step.1003d03.cons.ok
 55   AATGCTACGGCTCAACTTCGTCAAAACCATTCAAAGAGCCAAAGTGATTGG
      AATCATCAACAAACCAACAACAGAAAGATGCTTGGGATCCTGAACAAGAA
      ATCAACAAACAAAAAAGGACGATCCATATATATATTAATAGGTTTAGATT
      GATGTTTGACTATTAAATGTTGAATGCAGTCCATAAAAATACGCCCTATCA
      GTAAAAATAATTGATAAGGCGTATTTTTTATAAAGTAAGAAAGGGACTATA
```

```
        CAACTTGAAGTTAACTATATGAATAAATAGTGATAGAAAATGACTTTTTAA
        TTCGTCATTTCAAAATAGAGGATACATTTTCTTTGTTATTCATTTAATCAT
        AATTCTTGGAATGTTTGAATGATAAGATATACGTTAAGTCCACTTAAGACA
        ATTATGAGTAACATATCAAAAATATGTACACCAAAGTCCTTAGCATTTTGT
  5     TGTAAAGTTGTAATTTGTTTATTAGCAATTTCATCTTTTATATTAAAAATA
        TCAATTGTGGGAACGTTATGATCTAAAGTTGCAAAGGTTAGATCAGGTCTT
        CTGAGTTTACGATTTTGTATTCTAAGTCCTTCAAACGCTTGAGGAGAAGTG
        ACTTCATGAATGAGATGTAAATCAATGTATAATAATTGTGGTTCACCTTCT
        TTTCCATGAAGCACATGTTTTTTCCATACTTTATCAAACAGTGTTTGACCC
  10    ATAATTTTCTCCTCCTCTTATAGATATTTTTCTTTAAGCAATTTAAAAATT
        TCTGAAGTTCGATATTGTCCACCTAAATCTGCAGTTGTCTTATTAGATTGA
        ATAAACGAGTAAACAATTGACTCAAGTTCGTTAGCAGCATCATTTTGATTT
        AAACTTTCTCTTAAGCAAAGTGCTAAAGATAGAACCATACCAAATGGATTC
        GCTTTATCTTCATTAGCTATATCTGGTGCTGAACCATGAATTGGTTCATAA
  15    AGACGTGTACCTGTTTGACCAAAACTAGCTGATGGAGAAAGACCTAGAGAC
        CCTGGTATAACAGATGCTTCATCACTTAAAATATCTCCGAAAAGATTCTCT
        GTTACAATCACATCAAATTGCGTAGGTTGAGTGATCAGATGCATGCTACAA
        GCATCAACTAGCATATGATTAACTTCTACTTCTGGATAATCCTTTTTTACA
        TCGTTTACTATTTGTCGCCATAACTTACTAGAGGATAGAACGTTTTCCTTA
  20    TCAACAGAAGTTAATTTTTTACGTCTACGATTTGCAAGATTAAAAGCGATT
        CTCACAATACGTTCAATTTCTTGACTTGAATATTTTAATGAGTCTAAAGCT
        TCGGTTTTTTTAACATAGCTTGGTTCACCAAAATAAATCCCGCTCGTTAAT
        TCTCTTACTATAACTAAATCTGTGCCTTCTACAATATCTTGTTTTAGTGGT
        GAAAGATGACTAGCCCCTTTAGTTACAAATGTAGGTCGAATATTAGCAAAT
  25    AAATTTAAGGATTTCCTCAGTTTTAACAAACCGTGTTCTGGACGATTATTT
        GGGTCTGTCCATTTAGGCCCTCCAATCGCGCCTAATAAAATAGCATCTGCA
        TTTTTACATGATTGTAATGTTTCGTTAGTTAAAGGTGTACCATAATAATCT
        ATTGAAACACCTCCAAAATGATGACTTTCCAAATGGTATTCAAAATGATAT
        TTTTCACTTATTAATTTTAAAAGTTCTAATGTACCAGATAGAATCTCTGGT
  30    CCTATACCATCTCCAGGTAAGGCAACTATTTTATAACTCATGTACGTATAC
        CTTCTTTCAACTCATATTCTGAAATATATTTAGCATGAGCATCGATATATG
        CTTTACATGAAGCTTTTAATATATCGTGGTCTATACCTATGCCTGTCACTT
        CAATATGATTAATAATGATTCGTACATGTACTTCTGCTTGAGCATCAGTAC
        CTTCTGTTACAGAATCAATACGATAATCAATTAATTCTGCGTCTTTCTTGA
  35    AAAATTCGGTCAACAGCATTATAAATTGCAACAATTGAACCCGTTCCAATAC
        TTGAATCTTGTTTAACTTGTCCGTTTCTTTCCTTTATAACTACTACTGCAC
        TTTGTAGACCTTTAGATACGTATTGAAGTTGTAAGTTATCAAGTTGAAAAA
        TAGCATTGTGTTCATGTTCGGAGCCATGTATAATCGCATGAATATCTCTAT
        CGGATACATTTTTTTTCTTATCGGCAATTTCTTTAAATTGTTTAAATAATG
  40    TAACTTGATCTTCCAATTTAATTTCATATCCCAGAGCTTTAAGCTTTTCGG
        CAAATGCATGTTTACCAGACAATTTACCTAGTGGCAATTCTGTTGTATTCA
        CACCTACAAGTTGAGGTGTCATGATTTCATAGGTTTCACGATGTTTAAGGA
        CACCGTCTTGGTGAATTCCGGATTCATGACTAAATGCATTTTGACCGACTA
        TAGCTTTATTTCTAGGTACACGGATACCAGCATATCTTGAAATTAAGTCAG
  45    ATGTTTTCTTAGTTTCTTCAAGGTTAATTTGAGATTCAAGACCATAGTGGT
        CCTTCCTTACATATAAAGCCAAAGCGACTTCTTCAAGTGAGGCATTTCCTG
        CTCTTTCACCAATACCGTTCACGGTACCTTCAATACGTCTAGCTCCACCTT
        CAATAGCTGCTAAACTATTAGCAACTGCCATTCCGAGATCATCATGACAAT
        GTGCACTAAAGATAATTTTAGAATTTGACTTAACGGCCTGTGTTAATTGTT
  50    TAAAAATTTCGCCATACTCTGTAGGATAACTAAATCCAACTGTATCAGGGA
        TGTTGATAATTGTGGCTCCTGCGTTAATCGCAGTTTGAACACATTCAATTA
        AAAATGGAATTTCAGTTCTTGTTGCATCCTCTGGAGAGAATTGTACGACTT
        CAAAAAATTGTTTTGCATAAGAAACGTGTTCTTTTATTGATGTTAAAACTT
        CATCTTGAGTCATTTTTAATTTATGTTCTAAATGAATAGGGGAGGTTGCAA
  55    TGAATACATGTACTTGAGGTTTTACAGCTTCTTTAGTGGCTTCATATACAG
        CATCAATATCAGATTTTTTACATCTAGCTAAAC
```

Sequence 3451

step.1003d04.cons.ok
GAGAAGCAACAGTAATATCCACAACCTTTAGTGTTGTATCTATTACGTTCG
CTATTGTCATTGCCGAAACAATTAGAATGCAAGATCAATTTTTCTATTTTT
ATTTAACAGTTGTCATTTCATGCTTAATTGCAGCAATGATTATGCCAAGAA

5   TTTGGCCACTTAAAAATATTCCTGACGAATATGCTAAAGAAGTAAGTGAAG
AGGCTCGTAATGAACAGCTACCAGAAGGCAAAACAGCATTAAAATATGGTT
TTGATTTAGCAACTGAAGTTGGAATTAAATCGCCAGGGTTTAAAGAATTTT
TAATTTCAGGTTTTAAAACAGTTGTAGATATGTGGTTTGTAATTTTACCAG
TTGTTATGAGTATAGGAACAATAGCTACCATTATTGCTAACTACACGCCTG

10   TTTTTGAAATTATAGGAAAACCATTTGTTCCAGTACTAGAATTGTTACAAA
TTCCAGAAGCACATGAAGCATCACAAACAATTTTAATTGGGTTTGCCGATA
TGTTCTTACCTTCAATTCTTATTGAAGGGGTTCAAAATGATGTAACACGTT
TTGTAATTGGAGCATTGAGTATCTCACAACTTGTGTATTTATCTGAAGTAG
GCGGCGTGATTCTTGGTTCTAAAATTCCAGTTAGTATAAGTAAATTATTTA

15   TGATTTTTTTAATTCGTACTATCATTACGCTTCCAATAATTGCTTTATTAG
CGCATTTATTTATCGGATAAAAGATATACTCATTTTTAATTTAGTGAGAGT
GGGGGGAAGAAAATCATAACAACAGTAATATAAATTACTAATGGTTGAAAT
AAAAATTCTCATATATCAAACTTGTTCTAATTTTTATAATTATAAAAATTA
GACAATAAATCATATACGTATTCTAGCCCACTCCCTTTAGATAATTTATCT

20   TCTGTTCAAAGAAGATAAAACATTTGATTTAACAATTTCGCTTTTAATGAC
TGGTTATTTAACATTTATAGTCCTCCAATTATTTACTTAGACATTTATATT
ATATAGAAAATAGCGCTTTTTTACTATGATAGATGTACAATCAAAATTATT
ACAATTTTATAAATTCATTTAAAACTGCATATACGATAGCAATTCAGGGCA
AGAACTATTTTTTTAGGTAGCCCTCTTTATTTACTTGTCTTGCACGCGCTA

25   ATGCTAAACTATCTTCAGGGATATTATCTGTAATTGTGGATCCAGCTGCAA
TAAGAGTATGATTACCTACTGTTACAGGTGCTATAAGGTTCGTATTACATC
CTATAAAAGCATCTTTACCTACTATTGTTTTAAATTTGTTAGCGCCGTCAT
AGTTTACAGTGATAGAACCACATCCGATATTTGTACGCTCACCTATTTCAG
CATCTCCAATGTAGCTCAAGTGTGATACTTTAGCACCATCTTTAATATCTG

30   CTTTCTTCACTTCAACAAAGTTTCCTACTTTAACTTCAGAACCTAAATTAG
ATCCTGGGCGTAGTTGAGCGAAAGGTCCAACTGTTGTGTTCTCTCCAACAA
TAGAGTCATTGATAACCGATTGTTTTATGTTAGCATTCGAATGAATTGTGC
TATTATTAATTTCAGAGTATTGGCCTATCCACACATCTTCTTCAATCGTTG
TATGACCTCCGATGCGCACGCCCGGTTCAATAGTTGTATCTATTCCAATTT

35   TCACATCTGTTCCAATAAATGTTGAACTAGGATCAATGATTGTCACACCAT
TTTCCATATGATAGCGATTGATACGTTGTTGCAAAGCCTTTTCAGCTTCAC
TTAACATCAAACGATCATTAACACCAATGATTTCATCAAAATCTTCGGTAC
AATAAACTTCAGCTTTACCTCCGTCTTTTAAAATTAAAGACAAAACATCAG
GTAAATAATATTCTCCTTGAGCATTATCATTTTTAACTTGTTCTAACTTCT

40   CAAATAGTACTCGATTATTAAAGGCAAAAATACCTGAACTAATTTCTTTAA
TCGCACGTTCTGAGTCATTTGCGTCTTTCTCTTCAACAATACGCTCTAATA
TTCCATTATGATTTCTAATAATTCGTCCATAACCATAAGGATTGATAGTAG
AAGCAGATAATACAGTAACATGTGATTGTGTACTTTCATGATGTTCAATAA
GTGATTGTAAAGTTTGGTATGTAATAAGTGGTGTATCTCCACATACTACTA

45   GAGTAGTTCCTTCTTTATCTGCTAAATGTTCATGTGCCATTTTCACAGCAT
GAGCTGTTCCAAGTTGTTTATCCTGAAAACTATATAATGATTGATTACCCA
ATGTATCTTTCACACTCTCAGCGCCATGACCAATAATAGTTACAATTTGAT
CAACGCCAGCTTGTTTTTACGTTGTTAAGCACATGTTCAACCATTGGTTTGC
CAGCAACCTCATGGAGCACTTTATATTTCTTTGATTTCATTCTTGTGCCCT

50   TACCTGCTGCCAGAATAATCGCATGTCTTTGCATGAATGCTAACCCTCCAT
TAAAATTTACACTTCATAGTTATTATAGTTGAATAACAAAGAGAGTTTCAA
CTTATAAAAGGTTTTGTTAATCATTCATTAAAGGTTTGATAATTTTAGAAA
TTCAAATAAGTTGTTATGCTTTAATAAGGAAAACGAGGTCATCAAACCCAA
CAGTTACATTACATATTCAACCTCTTTATTAATCATTGGTATTTTATTTCA

55   GTGTCTAAATCATCATATTATCAATCATAATTTATGACGATAAAGTTATAA
TCATACAGACATATTAATTACGTCTATCATAAGTAAATTTTTATAGATGAA
TTGTTGTCAGAGAAACATATCAGTATTTTACTTCTATAATTCTTATTTATC
GTTAAATGTTTATTTCGTTTTCGTTTATGGAAATACAAAGTACATTAAAAA
TAACGGAGGGACTACAACTTCAATAAGAAGTCGCAGTCCCTCCCGAACAAT

Sequence 3452

step.1003d06.cons.ok

GTAGCCTAATGGCACATTATAATTATATTCAGAAATTTTGATTTTTCCATT
CTTATATACTTTTTCAACCACTGCTACGTGCCCATAATAGCCTCTCGTTGA
TTGCATAACAGCGTATTTTCGAGGCTTATGTCCAGTTGCATAACCAGATCG
ACGTGCATTGCTATACCAATTCTTAGCATTACCCCATCGATTTGATACAGG
TTTACCTAATTGAGCACGACGTTTAAATGCCCACCAAGTGCATTGTTTTTT
ACCATAATAATTATGTGCAGCATGGGAAATATGCGTTTGTTCAATAGCTTG
ATATGAGAGTGTTCCAGTAAATGTAATCATTGTAGCGATAATAATACGAGA
TAATAATTTTTTGAATTTCATTTCTTTCTCCCAAAAAATGTATTTATTTAA
TGACAGTAATAATAATAGTATTTAAAAATGCGCTATACAATATAAGAAATG
AATTATCAAAGTAAATGTAACATTAATGAAATATACATATTTTTATCGTAA
TAATCGTTTGAGTAAAAAAAGAGTGAGACTACATTCAAATAAAATGTAATC
TCACTCAGGTACGAGATGATTAACTCCATGTATATAAGTATGCGTAAATGT
TAAAGTCATGTCATAACATCAATTTATACTTTAGTGAATTGGTTGAAAACA
GCTAGTGCCCGTTTGATACTCGTCACTGGTATGTTTCCGGACATGGATGA
TTTCGAAAATATGTCAAGATTTCTTTTTGGCTCGTCACGTCTTTAGGGAAA
TTGATATCTTGATTAATCCAATCAACCAATTCACCTAATGGCGTGTCATCA
CCTAGGAAATTTTGCATAAATTCGTAAAAGCTCAATATCTCACCTCATAAT
ATATTTTATGAGCCGACTTTTATTTTAACATATTTTCGTTTTTTAAAATAC
TTGAAAAACTAAATAAAATGTTTTGAATTATAGTTTTTAAATATATTAATC
GCTAGAAAGGTCAATTTTGTTGTATAATTAAACTTACTATTGAATTATATC
AAAACTAGGGGGATATGCATGGATATAAAACACATGAAATACTTTGTTGAA
GTTGTCAAACAAGGTGGTATGACAAACGCTTCAAAATCCTTGTATATTGCG
CAACCAACTATCAGTAAAGCAATTAAGGATATTGAAGCAGAGATGGCTGTC
CCTTTATTTGACCGGAGTAAAAGAAGTTTAGTACTTACTGATGCAGGTAAA
ATTTTTTTTCAAGAAATGTCAAGAAATCATCGCACTATATGATAATTTGCCC
ACTGAAATTAATAGTTTGTATGGTTTAGAAACAGGTCATATCACTATTAGT
ATGTCTGCAGTGATGAGCATGCGTAAATTTATTGGCGTATTAGGAGACTTT
CATCAACTTTATCCGAATATTACGTACAACTTAATCGAAAGTGGTGGTAAG
ACGACTGAAAACCTTATACTTAATGATGAAGTGGATATTGGTGTGACAACA
TTGCCAGTAGATCATCAAAAATTTGAATGTATATCTTTAAACAAAGAAGAA
CTGACTGTAGTTTTAAATAAAGAACATCCTTTAGCACAAAAATCTTCTATT
AAAATGGAAGAATTAGCTGATGAGAACTTCATTTTATTTAATGAAGATTTC
TATCTCAACGATAAAATTATTGAAAATGCGAAGAATGCTGGATTCGTGCCG
AACATGGCCTCACAAATCTCACAATGGAATGTGATTGAAAATCTTGTCATT
AATCAATTAGGTATTTCCATATTGCCAGCCACTATAGCACAATTACTTAAT
GATGACGTCAAAATTGTACATTTGGAAAATGCACATACAACTTGGGAGCTT
GGTGTCGTTTGGAAAAAAGATAAACGTTTAAGTCATGCTACAAATAAATGG
ATAGAATTTTTTGAAAGAAAGATTATCCGAAGAATAATATAAATATAATAAC
AAACAATGCATTTAATAAAAGAGCGAAGTTGTTAAAGATGTGAAATAATAT
GTTTTTACAAAGTCATATTACAACGAAACGTAGGGTTTTGAATTATAAAT
TAATTCTGCCCTACGTTTTTTTTACTATCAATTAGTTCTCAGCACTATTAAT
TTGCTAGCGTTTTCATAGATTTAAGGAATATAAAATATCATGAATAAATAT
TTTTAATATATAATAGCAAACACTATAATAGAAAATGTCACTAAAAAAATA
CAGTAACAATTTGGACACATTTTTAGAAAGGGGTGTTGTTGATGGAAAAAG
CAAAATTTGTCATTAAGTTAATACTTCAACTTGCCCTTATCATGCTTATTA
CTTTTATAGGCACAGAAGTTCAAAAATTACTTCATATACCTCTAGCAGGTA
GTATCGTAGGGCTTATGCTTTTTTTTCCTATTGTTACAATTTAAAATTGTAC
CTGAATCATGGATTAATGTAGGAGCAGACTTTTTACTTAAAACAATGGTTT
TCTTCTTTATCCCATCAGTGGTAGGAATTATGGATGTTGCATCTAATATCA
CGATGAATTATATATTATTCTTTATTGTTATTATAATTGGTACATGCCTTG
TAGCACTATCATCAGGTTATATCGCTGAAAAAAATGCTAGAAAAAAGCAATA
CACGTAAAGGAACTGATCACTCATGAATGAATACTTACAAGCAGTTCTAAT
GATTTTGTTAACTATTGTTCTGTACTATGTTTCCAAAAAGATTCAAGATAA
ATACAATAATCCACTATTAAATCCAGCTCTTATTGCATCAATTGCAATAAT

TATTGTTTTATTGGTTTGCGGAGTAAGCTATAAGGGGTATATGAAAGGTGG
TACCTGGATTAACCATGTTTTAAACGCTACAGTTGTATGTCTTGCATACCC
ACTTTATCAAAATAAAAAGAAAATAAAAAAATATTTAACAATTATTTTCAC
AAGCGTGTTGACTGGTGTAGTTCTCAATTTTGTGTTAGTATTTACAACGTT
5   GAAAATCTTTGGTTATTCTAAAGACACAATTGTTACCCTGTTACCTA

Sequence 3453
step.1003d08.cons.ok
10   GTTATCCTCTTGTTAAAAAACATTTTCAATTTCTAATATTTTTAAAAATTC
ACTAGCTGCATATTTTAATGAAGATTCATCAATGTCAAAATGAGGATTATG
ATGAGGTGCAGTAATGCCTTTGTCTTTATTACCACAACCAGTTAAGAAGAA
TGCACCAGGTCGTACTTTAAGGTAATGTGAAAAGTCCTCACCAATCATCAT
TAAGTCTGACTCATTAAATCTTAAATGTAAATCATTAGCTGCTTGTTTGAC
15   TACATCATATGATTGTTGATGATTGTGTACTGGTAAATAACCTTTAATATA
TTCCAATTCATATGTAATATCGTTTGATAAAGATAAACCTTGTAAGAGCTT
ATCCATTTTATTTTGAACATGACTTTGTAATTTTGTGTCAAAAGTACGTAC
AGTACCTTTACAAAAAGCTGTATCTGGAATAACACTATCTGTTGATCCTGC
TTGAATCATTCCGAAAGTAAGAACAGCTTCCTTTACTGGATCAATTGTTCG
20   AGAAATTATTTTTTGGGCACTTAAAATAAACTCAGCCATAATGACAATAGG
GTCTATTGTTTCGTGTGGTTTTGCACCGTGACCACCTTTTCCATATATAGT
CACACTAAATTCATCTGGTGAAGCCATTATTGCTCCAGGTCTAGAATAGAT
TGTCCCAGATGGATAACCACTCCATAAGTGTGTGCCATATATTTTATCGAC
ATTCTGTAGACAGCCATCATCAATCATCTCTTGAGAACCACCTGGCATAAT
25   TTCCTCACCATATTGGAATATAAAAACAACATTACCTTTCAATAAATGACG
ATGCTCATGAACAATTTCAGCTACTCCAAGCAAAATAGCTGTATGTCCGTC
ATGTCCACAAGCATGCATGCACCCTTTATTTTTTGAACGATAGGGTACATC
ATTTAATTCTTGAACAGGTAGTGCATCGAAATCTGCTCGTAATGCAATCGT
TGGTCCATTTGAATCAGATCCTTTAAAAGTTGCTTTTATACCATTACGTCC
30   AACTGGTGTTTCTATGGTGCATGCTAATTGGCTCAATTGATTTACAATAAA
ATCATGCGTACGTTTTTCCTCAAAAGAAAGTTCTGGGTATTGATGCAAATA
TCGTCGTAATTGCACCATTCTCTTTTCTTTATTACTTGCTAATTGAAACCA
ATCAAACACATGTATCACTTCTTCCATCTTAATTATTGATATTAGTATATC
ATTATTGTTTTAGCTTAAATAGACTTTGTTTTGTTATTAAGATATCACACG
35   TTTTTTACACTATCTATTAAGATTTGGAAACATTAAAATGATTAAATATAA
ATGAATAGCCCCATAAAAATACAACACAAATCATAATACCTCACTTAAAAT
TAATAAGAAAAAATCGTCCATAAAATTTTTGTTATTAACTGTTATTATAGA
AGATAAAGTTTATGTAAAAGTATTCAAAAATATGTAGAATACTAAAATACT
AGGTAGTGATGTGAATTAGGCGTATTTCTAAGAATTTAATTGGTATACGAC
40   ATTTTAATATTATATGAAGTTCTGGCTAAAGGAATATTAGAATTAGTAATT
ATAGTTTATTAGAAGTGGAAAAAAATTAAAAAGTAAGAAAATTATTATAAAA
TAATACCTTAACTTATATGTAGAAATATAAAATTATTATGATATAGTTTGT
TATAATTTTTGATGTTCACAGTTAATCACATTAAAATAGAGTTATCAATAA
TATAAGTTATGGGGTATTTAAGTGGAGGAAATAAGACCAATAACCTATCAA
45   GATAAAGAGGCATACTATTATTATATTCAAGAGTGGTATGAAAATGAGGAA
AAAGTGGTACCAGGGAATACAGATATTGCCAATTATAGTTCATTTAACAAT
ATGGTTGATCGGCTTAATTGTAGTGAAGTTGATGAGGGTTTTGTACCGACT
ACAACACTATTTTATTTTAAAGATTCAATTATTATAGGTGCTGTTGATATC
AGACATCAATTAAATGATAAACTATCTAATATTGGTGGTCATGTGGGATAC
50   GGTGTAGCTAAATCTTATAGAGGGAAAGGTTACGCTACTATCCTTCTAGAA
AAGGCTTTAGATGAACTTAAGACATTAAATGTAGAGGTCGTACTTATGACT
TGTAATCCACTTAATTTTGCTTCTCAAACAGTGATGAAGAAATGTGGTGGA
TATCAAATTGAATCCTATATTAAAAAAAAATGGTAAACCTGTTCATCGATAT
CATATACCTAATACAAAATAATCATGATTATTAAGAATATATTGTGATAAA
55   GCAAATATATGATTTTTATTTTTAAAATGATTAGAAGAAAAGAGCTTTTTA
AAGTCAATATTATTTGAGATAGGAGCTTAGACTTTTAAGTTTATTTAAGCA
AAAGAATTTTAATGGAGACTTACGATTTAACACATCTTGTTATATTAAACT
TTGTGAAAAACTTAACCATCTAAATAATCCTATATAATACATTAAACGCGT
GGGGGGGGATTCATAATAATGAAATATATAGATGAGCAAACACAAGCTCAAT

974

TATTAGACATGAATGAAGTCATTTTAGAAGTAGAAAAAGCTTTGCAAGCTT
TTTCAGAGAATAAGACGATAACGCCATTAAGATATGTTTTGCCATTTAATG
AGCAAAATCGTTATTTAGTGATGCCAGCATTATCAGATGAATTAAATATCG
TTGGACTTAAAACAGTCTCATTTGCACCTGAAAATTCAAAAAAAGGGAAAG
CGACTATTACTGGATCAGTTATTTTAAGTGACTATGAAACAGGAGAAACAT
TGTCTATATTAGATGGTGGTTTTCTTACTAAAGTAAGAACTGGTGCAATTT
CAGGTGTAGCTACTAAATATCTAGCAAAAGAAACGCTAAAACACTTAGTG
TAATAGGGGCAGGTGTACAAGCTGAAGGTTTAATTGAAGCGATACTTGCTG
TTAGAGATATTGAAAAAATTCACATCGCAAGTAGAACGTTCGAAAAAGCAG
AAAAATTTGCTCAAAATATACGAAATCGATTTAATATTAAAGTGAGTGTAT
TTAGATCGGCAGATGAAGCGATAGACAGTGCAGATATTGTAGTTACAGCAA
CAAATGCAAATCAGCCCGTTTATACTCATTCTTTACATCCAGGCGTGCATT
TAAATGCAGTCGGATCCTTTAAACCAGATATGCAAGAAATACCTTCAGAAA
CAATGCTTGTTGCTAATAAAATTGTTGTTGAATCTATGGAAGCAGCTTTAG
AAGAAACAGGTGATTTAAAAATTCCTCAAGCAGAAGGAATATTAACTAAAA
ATATGCTACATAGTGAATTAGGCGACATTATTTCT

Sequence 3454
step.1003d09.cons.ok
GTGAGGGTGCTTCTGGTAAAGATGCTATCACTAAAGCGCATGAGCTAAAAC
CAGATTTAATATTAATGGATTTACTTATGGACGATATGGATGGTGTAGAAG
CAACTACTGAAATAAAAAAAGATTTACCTCAAATTAAAGTAGTCATGTTAA
CAAGCTTTATAGAGGATAAAGAAGTTTATCGTGCACTTGATTCTGGAGTAG
ATAGTTATATTTTAAAGACAACAAGTGCAAGTGATATAGCTGACGCTGTGC
GTAAAACGTATGAAGGTGAATCAGTATTTGAACCAGAAGTGTTAGTAAAAA
TGCGCAATCGTATGAAAAAACGTGCCGAGCTTTATGAAATGTTGACAGAAA
GAGAAATGGAGATCCTATTACTTATAGCTAAAGGATACTCTAACCAAGAGA
TTGCAAGCGCCTCTCATATCACCATCAAAACAGTAAAAACTCATGTAAGTA
ACATACTAAGTAAATTAGAAGTACAAGATCGAACACAAGCAGTAATATATG
CGTTCCAGCATAATTTAATTCAATAAATTATTAAAGGCGAAAGTAAAGATA
CATCTATCGATACTTTCGCCTTTCAATATTATCAGTCGCTTGTGTGATGTT
CTTCTTTAGATTGATGATTGATATTATACGTGTCATCTTCATTAGAGATAT
CTACATCTTTGCTATCTTCATAGTATTCATAAGTCGTATTTTCGTTATAGT
GATTTTCATTGTTATCATCATGTTTAATAGCGGCTTCTTCAGGCGTGTGAC
CAGCTATGACTTTTCTTTGGTGAATAATTGCATTGATTTCAGCACCGATAA
TAATAATAAAGCTTGTGATATACAACCATAAGAATAAAATGATAATACCAG
CTAAACTTCCGTATGTTTTCGAATAGTTACTAAAGTTTGAAATATAGTAAC
CGAATGCAAAGGACCCTAGTAACCAAATAATGGAAGTGAAAATAGCGCCAG
GAATAACAGAACGTAATTTTGTTTTAACATTAGGTGCAACTGAATAAAGTA
CAGTAAATATGATGAAAATAATAATCAATGGAATAACAATTCGTACTAAAT
TAAAAATCCATTCAATTTGATTATCAATACCTAGTGGTCCAAATAAAAATT
TATTAATGACTGGACCTAGTGTTATAAGTACTACAGCAACAACAAATACTG
CACCTAAAACAAGTGTATATAGAATACTTAATAATTTAACTACTACACCGT
TTCGAGAATCCTCAACGTCATAAGCAACATTAAATGAGTTAATAATAGCGG
ACATTCCATTTGATGCTGACCAAATAGCTAAAATTAACCCAACTGACAAAA
TACCTCCACTGGCAGTGTCAGAGATATCTTTAACAATTCCACCAACGATGC
TAGCTGTTTCTTGATCAGGGACATATTGACTTACTTTTTGATTAATCTGAT
TAGCATCAATCGTGATAAATTGACCAAGTAGCGTAAGTAAAAAAATTAGCA
TTGGGAAAAGTGCTAATACAAAATGATATGTCATCTGTGCTGCTAAACCTG
CAGCGTCATCTTTACCTATCCTATAAATCAGGTAAGAAAAGAAATTAGAGT
TTTCGGTGTATTTTGCTGGTTTATTCAGCCGGGAAACAAAAAATACTTGAT
TCTTTTTTTTAGGTTTCTTAGATTGGAATTCTTGAGGTTCTATATATGTAC
GGTCAACTTCTATTTTTTTATTCTTTTTATGCTCTTTATCTTCTATTGAAT
TAAGATATTTAGAAGTTGTTTTCTCTTTTTTTTGACATAACAATCTCCTTTG
CTTTTTATGTTTTACTTTTAAAAATTGATATTCCCGTATTTGCGGTATAGTA
AAACAAATATGGTTTTATTTATATATGATAAATCCAATACCCTTTAAAAAG
TAATGTATTTTTAACACTCAATGATTATTTTGGCCTAAATTAAATAGAAAG
GCAATAGTTTTTAAATATTGATTTATAGTTTTTAACATGAGTTAAAGTATT

AAACAAGTGCGAGTCAAAATTTAACAATTATTGTAAAGTTTATAATCAATA
AAGATAAAAAGAGTGGAGCAGAATCCATCTAAAATGAATTCCGTTCCACTT
CATAGTTATTAAATATCATAAACTTATATGCGTACATTTTCCAATTGATGT
CCTACTTTCTCTGAAAATGCTACGCATTTGAAAGTTTATTTACCTAAGCGT
TGATTTTTACGATTTACAAAAGTTTCTTTGGCATCTTTAATAGAACGTTCT
AATTCAGGATTATTACGACGTATTTCATCTAATGTATCTTTCCAATACATT
ACTTCATCTTTGATGCTACTTATTTTTGAAGGTTTACGTGTACGATTTCCT
TCTTTGATGTCTTTGAATGATTGTTTTAAAGAATGACGTGTAGATTTATCA
GCAAGAGCAGCCGCGCCACCAATTACGGCACCAATTAAAATACCAGGTACT
AATTTATTTTCCATAGTTGAATTACCTCTCTTTCAAATTTGCATCTTTTAC
GATGTAGTCTATTAAATTATCACAAGATGATTGCACCATCTCGAAAACACC
TTCAAAATTATTTGTGTAGTATGGATCTGGTACATCACTCTCTTCCATGTT
ACTAAATTCTAGCAATTTGAACAATTGTCCTTGTAAATTTGGATTGATTTG
TTTGATATTGTCTACGTTACTTTGGTCCATAGCAATAATATAGTCAAAATC
ATCATCAGGTTCGAAAAGTTCACTCACCATACCATCATAAGGAATATGGTA
CTTCTGTAGAATTTTTTTGTGTTCCGTTATGTGGAGGTTCGCCTAAATTCCA
ACGTCCTGTTCCTCTAGAATGAACTTTTATATCTGAAATACCTCTTTCTTG
TAGTCTTTGTCTCATGATAGCCTCAGCCATTGGAGAACGACATATATTACC
GAGACATACAAATGCTACATGTATCATTAGTATCACTCCAAACTATGTATT
TATATACAAATAGCTCATATTATTACAAAGTTCAAACTTTTATACCTAAAA
ATTATATTGAAGATTTCAAAGACAGTTTAAACGAACTTGGATTAGCTAAAG
AAGGTAATTACGAATATTGTTTCTTTGAAAGTGAAGGTAAAGGGCAATTTA
AACCAGTAGGTGATGCAAGTCCTTATATAGGGAAGTTAGATAGTATCGAAT
ATGTTGATGAAATAAAACTTGAGTTTATGATAAA

Sequence 3455
step.1003d10.cons.ok
GTATCTAAGCCTAAAGCTAATGAAGCAGTAGTGACGAACGAGTCAACTAAA
CCAAAAACAACAGAAGCACCAACTGTTAATGAGGAATCAATAGCTGAAACA
CCCAAAACCTCAACTACACAACAAGATTCGACTGAGAAGAATAATCCATCT
TTAAAAGATAATTTAAATTCATCCTCAACGACATCTAAAGAAAGTAAAACA
GACGAACATTCTACTAAGCAAGCTCAAATGTCTACTAATAAATCAAATTTA
GACACAAATGACTCTCCAACTCAAAGTGAGAAAACTTCATCACAAGCAAAT
AACGACAGTACAGACAATCAGTCAGCACCTTCTAAACAATTAGATTCAAAA
CCATCAGAACAAAAAGTATATAAAACAAAATTTAATGATGAACCTACTCAA
GATGTTGAACACACGACAACTAAATTAAAAACACCTTCTATTTCAACAGAT
AGTTCAGTCAATGATAAGCAAGATTACACACGAAGTGCTGTAGCTAGTTTA
GGTGTTGATTCTAATGAAACAGAAGCAATTACAAATGCAGTTAGAGATAAT
TTAGATTTAAAAGCTGCATCTAGAGAACAAATCAATGAAGCAATCATTGCT
GAAGCACTAAAAAAAGACTTTTCTAACCCTGATTATGGTGTCGATACGCCA
TTAGCTCTGAACACATCTCAATCAAAAAAATTCACCACATAAGAGTGCAAGT
CCACGCATGAATTTAATGAGTTTAGCTGCTGAGCCTAATAGTGGTAAAAAT
GTGAATGATAAAGTTAAAATCACAAACCCTACGCTTTCACTTAATAAGAGT
AATAATCACGCTAATAACGTAATATGGCCAACAAGTAACGAACAATTTAAT
TTAAAAGCAAATTATGAATTAGATGACAGCATAAAAGAGGGAGATACTTTT
ACTATTAAGTATGGTCAGTATATTAGACCGGGTGGTTTAGAACTTCCTGCA
ATAAAAACTCAACTACGTAGTAAGGATGGCTCTATTGTAGCTAATGGTGTA
TATGATAAAACTACAAATACGACGACTTATACATTTACTAACTATGTTGAT
CAATATCAAAATATTACAGGTAGTTTTGATTTAATTGCGACGCCTAAGAGG
GAAACAGCAATTAAGGATAATCAGAATTATCCTATGGAAGTGACGATTGCT
AACGAAGTAGTCAAAAAAGACTTCATTGTGGATTATGGTAATAAAAAGGAC
AATACAACTACAGCAGCGGTAGCAAATGTGGATAATGTAAATAATAAACAT
AACGAAGTTGTTTATCTAAACCAAAATAACCAAAATCCTAAATATGCTAAA
TATTTCTCAACAGTAAAAAAATGGTAAATTTATACCAGGTGAAGTGAAAGTT
TACGAAGTGACGGATACCAATGCGATGGTAGATAGCTTCAATCCTGATTTA
AATAGTTCTAATGTAAAAGATGTGACAAGTCAATTTACACCTAAAGTAAGT
GCAGATGGTACTAGAGTTGATATCAATTTTGCTAGAAGTATGGCAAATGGT
AAAAAGTATATTGTAACTCAAGCAGTGAGACCAACGGGAACTGGAAATGTT

```
     TATACCGAATATTGGTTAACAAGAGATGGTACTACCAATACAAATGATTTT
     TATCGTGGAACGAAGTCTACAACGGTGACTTATCTCAATGGTTCTTCAACA
     GCACAGGGGGATAATCCTACATATAGTCTAGGTGACTATGTATGGTTAGAT
     AAAAATAAAAACGGTGTTCAAGATGATGATGAGAAAGGTTTAGCAGGTGTT
5    TATGTTACTCTTAAAGACAGTAACAATAGAGAATTACAACGTGTAACTACT
     GATCAATCTGGACATTATCAATTTGATAATTTACAAAATGGAACGTACACA
     GTCGAGTTTGCGATTCCTGATAATTATACGCCATCTCCCGCAAATAATTCT
     ACAAATGATGCAATAGATTCAGATGGTGAACGTGATGGTACACGTAAAGTA
     GTTGTTGCCAAAGGAACAATTAATAATGCTGATAATATGACTGTAGATACT
10   GGCTTTTATTTAACTCCTAAATACAATGTCGGAGATTATGTATGGGAAGAT
     ACAAATAAAGATGGTATCCAAGATGACAATGAAAAGGGAATTTCAAATGTC
     AAAGTGACGTTAAAAAATAAAAATGGAGATACCATTGGGACAACGACAACA
     GATTCAAATGGTAAATATGAATTCACAGGTTTAGAGAACGGGGATTACACA
     ATAGAATTTGAGACGCCGGAAGGCTACACACCGACTAAACAAAACTCGGGA
15   AGTGACGAAGGTAAAGATTCAAATGGTACGAAAACAACAGTCACAGTCAAA
     GATGCAGATAATAAAACAATAGACTCAGGTTTCTACAAGCCAATATATAAC
     TTAGGTGACTATGTATGGGAAGATACAAATAAAGATGGTATTCAAGACGAC
     AGTGAAAAAGGGATTTCTGGTGTTAAAGTGACGTTAAAAGATAAAAATGGA
     AATGCCATTGGGACAACGACAACAGACGCAAGTGGTCATTATCAATTTAAA
20   GGATTAGAAAATGGAAGCTACACAGTTGAGTTTGAGACACCATCAGGTTAT
     ACACCGACAAAAGCGAATTCAGGTCAAGATATAACTGTAGATTCCAACGGT
     ATAACAACAACAGGTATCATTAACGGAGCTGATAATCTCACAATTGATAGT
     GGTTTCTACAAAACACCAAAATATAGTGTCGGAGATTATGTATGGGAAGAT
     ACAAATAAAGATGGTATCCAAGATGACAATGAAAAAGGAATTTCTGGTGTT
25   AAAGTAACGTTAAAGGATGAAAAAGGAAATATAATTAGCACTACAACAACT
     GATGAAAATGGGAAGTATCAATTTGATAATTTAGATAGTGGTAATTACATT
     ATTCATTTTGAGAAACCGGAAGGCATGACTCAAACTACAGCAAATTCTGGA
     AATGATGATGAAAAAGATGCTGATGGGGAAGATGTTCGTGTAACGATTACT
     GATCATGATGACTTTAGTATAGATAATGGTTATTTTGACGATGATTCAGAC
30   AGTGACTCAGACGCAGATAGTGATTCAGACTCCGACAGTGACTCGGACGCA
     GACAGCGATTCTGACGCAGAC
```

```
     Sequence 3456
35   step.1003d11.cons.ok
     GCATTTGCTAAGGGACTTGTTTCATCGTTTGGGCCATTTCCTTCAGCAACA
     ATACAACATATAGGCTATGGCGCCGGCAGTGAAGGATTTTGATTTCCCTAAT
     ATATTAAGGGTTATTCAATTTGAATCTGAATTCGAGCAACAAGATAGCGTC
     CAAGTAATAGAGTGTCAAATAGATGATATGACACCTGAAGCATTAGGTTAT
40   TTTATGAATAATGCGTTAGAGCAAGGTGCTTTAGATGCTTACTATACGCCT
     ATATTTATGAAAAAAAGTCGCCCAAGCACGCAGTTAACGTTAATATGTAAA
     TTACATGATAAGACATATTTCGAACAACTTATCTTACAAGAAACAAGTTCT
     TTAGGCGTCAGAAGTACTTCTGTTAATAGAAAGACCTTGAACCGCGCATTC
     AAAATTCTTTCTACACAACACGGCACTGTTTCCATTAAATTTGGCCTACAA
45   AATGGAAAAATTATGAAAATGAAACCCGAGTATGAAGATTTGAAGAAAATA
     GCTAAAACTACAAAACAACCGTTTCAAGTAATTCATAACGAGGTATTACAA
     CAACTCTATCAAACATATCATATAGGAAATATACTTCAATAAATTAGTTTT
     ATTCATTTAAATTTTGCACATCAAAAAAGGATAGGGACAGTCATTCTACAT
     AAAACAGTCACCTATCCTTTTATTCTATTCAAATGCTCGATATAGTTCTTC
50   ACTAGAATTATTCCACATTTCAACGTTATGATTTTGTAGAAATTCTTTTAA
     AATTTGTTTGTTATCATCACCGATTTGGTCGATGATAAATTGATGTTTCAT
     AGACTTATCCATCATATTAACATGTTCTGGCATGCTCTTATAACCACGTCT
     AATACTTTTGTTAACTAATAAATAGCATGCAGTCACACCCGCATAATATGG
     CCCTTCTTCACTACGTTCGGTCGTCACCCAACATAGCCAGTATGCTTTAGC
55   CTCCTCACCTTCAACTTGTGATTTGTCAGTAATCCACTTGACACCCTTTTC
     AACTTCAGCACGTGCGTGCATACCACCAATATCAACAAATGCCTTCTGATT
     TTCAACATCTATAAAAACCGGGGCAATATTATCTAAACTAATCGAACCAAT
     ATTGGTTCCTTTATGACCATCTAGTGGATCATTTTTTATAATATTGAAATT
     GAAACCTTTTTTCTCAGACAAACGAGATACCTCCTAATAAGTCATCAAATA
```

977

AAGACTATATATCTTATTCTTAACTGTTAATTTAATTCATTAAGTAAGTTA
ATAATAGAACCTTTAATTTGATAATCTTTAATTTTTAAAATCATCTCTTTA
GGATAAATATAATTTATGGTCTTCTCTATTAATCCCTGTAATACTATTAATC
ACCAATGATTGATTACTTATTTCTGTAATTCCACCATTACTTTTTAATAAA
5   TGAATCGGTTTGCGATTTGAGCCTGGGCGATCATAATCATAAGGTAAATCT
GAAAATGCTTCACTTACAAAGTAATAATCAGGATTAATACCACCCGCTTCA
AATAATTCTTGCAATTCCGAAATGGTAATAATTGAACCGTCGAAAGGAATA
TATTTAAATAAATCTCGATTGATAAACCGTCTTGATAAATCACTTAATATT
GTATCATTTTCATGAATCCATTTCTTCAAGTAATACAATACAACTGCTTCA
10   TCAAGTTCTACATATTGTTCAATCGTCATTGTTCCTTCAAAGAATGGTATA
AAGTCTTTTGGATACATTTTAAATTCATATCCTTCATTATAAAGCTGCTTA
GCTCGTTTTAAACAATTGTTTAATAAAACTTCCCCACCACGGCTTACTGGA
TGAAAATATATTTGCCAATACATTTGATAACGACTCATAATAAAATTTTCG
ACAGCATGCATACCACTATCTTTAATTAACACTTCTTCTTTAGATGGTCTC
15   ATCAACCTTAAAATACGCTCCATATCAAATGAGCCATACGTTACGCCTGTA
AAATATGCATCTCTTTGTAAATAGTCCATTCGATCAGCATCAATTTGTGAG
GAAATCATAGAAATAACAAGCTTATTATCGTGCGTTTTATTAATTACATCA
GCTACTTGTTTAGGAAAGTTATCAGAGACTCGACTTAGTACACCATTAACC
TCAGTGTCTCCTGTGATAATAGCTTGTGTAAATGCTTCATGGTCTGTATTA
20   AATATTTTTTCAAAACTATGAGAGAAAGGACCATGACCTAAATCATGTAAC
AACGCAGCACACAACGCAAGTGGCCGATCAGTATTATCCCACGCGTCTTGT
CCTATAAATGTCTCATCAATCATTCGGCGCACAATTTCATAAACACCCAAA
GAGTGTCCAAAACGACTATGTTCGGCTGTATGAAATGATAAATACAAAGTA
CCTAGTTGTTTAATACGTCTCAAGCGTTGGAATTCTTTTGTTTTAATTAAG
25   TCCCAAATCAATTGATCCTGTACGTGGATAAATCGATGAATTGGATCTTTA
AATACCTTTTCTTCTGTTAATTTACTCGTTGCATATGATGATAGCGTCATA
CTTTCCCTCCTATCCTATATATTCCAAAAAACATTTATTTTTTGTTCCTCA
TATTACTAGACGAAACAATGTTAATTATTCATATGTTTGTTAAAATGATTT
CTTCATTAAAGCTAAATCCATGATTTCACCATACATTTCAGGTTGATATGA
30   GCGAATAATCTCGAATCCATGATTTTGATAATATTCGATGCCATTACTATT
ATTGTTATCAACTTCTAAATAAACTGTTTCATACTGGTCTTTAAACCTTTT
CAACCCCGCTTCTAGCAAACGAGTTCCATAACCTCTATGTTGAGATTCAGG
TCTAACATAGTGTGCAGATAAATAAAGCTCTTCACCATAAATGAAATTTGC
AAAACCAACAATCTCACTATCTTCTTCTGCCACTAAAAATAATTGCTCATT
35   TAGTCTCTTTTTTAAATGCGTTTCATTATATGAAGCGGCCAACAGTTCATT
AACTGTTGAAGCAGCATAAATATTTAAATATGTGTTATACCAAGCTTTTGT
CGCAACATCTCTTATCGCTACAACATCTTTTTCAGTTGCTTGTCTGACACT
ATACATGTGACTTCTCCCCTAACTATTTACTTCACTATTAATCGATGGGTG
GATTGGTAACTTGTGCAAATAATTGTTTAAAATAATTAAATAGTGACTCAG
40   GCTTATCATTAAGAACTGCAATAGGTGCATCATATCCCATCGCTCCGATGG
GATCTTTTTTATTGATGACTAAATCTGTCATATATTTATTCTTTAGAGCAC
AGTGGCGATGATATC

45   Sequence 3457
        step.1003d12.cons.ok
        ATTTTTCAATCATCGATAAACTTAAACTAATTCATTCTCATATTGGGAGAT
        CTTAATATTTTAATTCATATATAAATAGTAATTTTAGACACTGATCAATCT
        TTTAAGTTGGTTATTTTAAATATCAGGTTCCAATTTTTTCTTTTACAACTT
50   AACAATTCTACTTACCTCATATACCTTCTAAAACAGGTTTCATAATTAATG
        ATAAATGACGTTTGTATCACATCAAAGGCATTTGCCTATATTTTTCAAAAA
        ATGTGTCCTTTTTTTTTCAATCTATTTCGATATATATAGTGAAAGGAGGTTT
        TATTATGAATGAAGTCAATCAAGTGACGATTATCTTTACTCAAATCAAACA
        TAATTCTGAAGGTAAAGTAGTCGAACATAAGCGTCGTTTTTCTAACATTAA
55   TCCGCTTGTTTCTAATGATCAAATCAAATCTTTTCGTTCCATCATTGAACG
        TATCAGTGGCGAAGAGTATGACAAAGTCGAGATCGTTAAAACAGAATCTTT
        AATTTAAGGAGGATACATCAGTATGTCAAAAACTTTGGAACTCGTATTTAA
        ATCAAATCTTAATAAACCCGTCAAATTACTATTACCCGATTTTAACGCTAT
        TACTACAGAACAATTAATTAAAGAAAGTATGAACCAACTTTTAGAATTAGA

978

```
CATTTTAAGATTTAGTTTGGGAAAGCCTGTCAAAATTTATGCTGCACAACT
TATTGATAAAAGTACAACTGTTATATTTGAAGATAAAAATTAATTCATAAA
ACATTATTTTTGACGCTTTATTTAAAGTACTAAAATCATAAAAAAGACTTA
TCGAAGAACAAATTTTTTAAAATTGTTCAAGTCTAAAAAAGAACTTAATTT
ACATTGTTCTTTTATTAATATGAAGTGAGATATCAAATTCAAAATGATGTC
TCGCTTTTTTTAATTCTAAAACTTGTTTAATGTGCATTTTATTATAATGAT
TCATCAAAATAGTATCGGCTGTATGACACCTCGGAACATTCACAACTTATT
TCTATACTTTGCTTATGGCTCAATCATTTTAGAAACATGATTATGAATGGT
ATGATAATCATATTAAATTGAGAAAGTTGGTTTTTATCATTGTTTTTTCTA
TATCTTTTAGGAATTTTTGTTGGCATGTTACTTCCTATTCAAACTTCAATC
AACTCAAGGCTAAGCCAATTTACCCGTTCATCTTTTTATGCTTCAACCATA
TCTTTTGCTGTCGGAACAATATGCTTACTTGTGCTTAATATCATTATTCAT
CCACAAGTATTAACACCAGAATTCTTTTCTAAGCAAACACTTAACTATACT
TGGGTCCTAGGAGGACTATTAGGTGTTATCTATTTAACTGGAAACTTATTA
TTATTACCAAGATTAGGCGCAGCACTTACTGTTGTTATTACGGTTACAGGG
CAAATTATTATGGGTGTAATCATTGATACATTTGGATTATTGGGCGCCCAT
CAACAATCTTTTACATTATTTAAAGGTGTTGGAATTATATTTTTAATTACT
GGAATTATATTTATGAATTATGTTCGTAGACATCCTGTTAATAGACATAAG
AATACACCAATAGTATTCTGGTTGCTTATTGGATTCGTATTTGGTTTCGCT
CCGCCAATCCAAACAACAATCAACAGCACACTGGCTCAACACACACACTCA
TCCATATTTGCTTCACTCATATCATTTAGCGTTGGTACAATAGCGTTATTT
ATATTAACATTAGTATTTAATCGTAGTCTGAAAATATCATCTACTCATAAA
ACACTAGGTAAGATTAAATCTATTTATTTTATTGGAGGGATACTAGGAATG
GCATTCGTAACATCTAATATTATTTTAATGCCTTTCTTAGGAGCAGCTTTA
ACTACTATAATTGCTATGATGGGCCAAATGATTATGGGTATTATTATCGAT
CATTTCGGTTTGTTGGGGTTCACCTAAAAACAAAATAACATCACGTAAGTTA
GGTGGATTATTATGTATAGCTATAGGTATTATTTTATTACGATTATTTTAA
ATTCCTTATTCTGAATAAAAATTATCCGCTTATATTAATGTAGATTCCTCA
TACAAATAAATGTCTATTTCGAACTGCCAACAAAGATTTTGTCTTTGTTGG
CAGTTTTTATGAGTGATTCTTCTTATAGTAACATTTATGTAGAATCATATA
TTTTAAAATTAAAGATAGCCTTCTATTTTTCAATTTATTAAATTGATTTAC
TTTATTTTTTGAAATAAAACAATAGTAATTGTGCACCATATTGTTCCGTCT
CTTCTCCTGTAGTAGCTAATAAACGTATCGTACCTTGCTTTAATGCGTTTG
ATTGTTTTAATTCAATCATCGCGATGACCATCGCCATAAGACCACCTTTCA
TGTCAGTGGTACCTCTGCCGAATAACTTGCCATCTTTATCTGTGAGTTCAA
ATGGAGGAAAAGTCCAATCATCATGGTCACCAGCATCTACAACATCCATGT
GTCCACTGATTGCTAAGACTGGAGCGCCACTACCAATTTCAGCAACTAAAT
TAGCGCGTGAATCATTAACTTTAACAATCTTAGAATCAATATCATATTGAC
TTAACAGGTCTTTTAAGTACTCACATACTTTAATTTCATGATCATTCTCAG
TTTGAATCTTGACGATATCAGCTAATAATCTTATTTTGTCTTGCTCACTCA
ATACAGTCATTTCACAGTCCTCCTCATTTTTAGTAATGATTTTCACTTTTT
AATAAAAAATAAAAATCCTTGTACACTTAATATTATCAAAAACTTTTATTA
ATTAAAAATTATATGCTAAAATAAATGGATTTTTAATAAAAAAATAAAATATT
AACTTAACATTTCTTATAAAAAACAACAATGTATAGAAATATTAAAAATTT
TTCATTATTTATAGTTTGAATCGTGTTATAAAAATGAAAAGCTATGAGATTA
AGATGTAATAATACACACATATATAAATATAACTTCGTGCCACGCAGACGG
TTAAAATCATGTATAATAGCTAATAATACATATTAGGAGATACATAATTAA
TGATGAAAATAAATTAACATTAAAAGAGAATCTATTTATCGGCTCAATGC
TGTTTGGTCTTTTTTTTGGTGCTGGAAATCTCATTTTTCCAATTCACTTAG
GTCAAACTGCGGGGGCAAATGTATGGACCGCCAATTTAGGATTTCTTATCA
CGGCTATCGGACTACCTTTTTTAGGAATTATAGCGATAGGTGTATCTAAAA
CAAACGGGGTCTTTGAAATTTCCTCAAGGATAAGTAAAATATATGGTTATT
TGTTCACAATTGGCTTGTATCTTGTTATAGGTCCGTTTTTTGCGTTGCCAA
GACTTGCGACGACGTCATTTGAAATAGCATTTTCACCATTTATTTCATCTG
GTACGGCCCAAGCGTTGTTGCCTATTTTTAGTATTTTATTCTTCGGAGTAG
CGTGGTTATTTTCGCGTAAACCTTCTAAAATATTAGACTATATTGGAAAAT
TCCTTTAGAGCACAGTGGCGATGATAT
```

Sequence 3458
step.1003e02.cons.ok
GTCTCGCAGTCAAGCTCCCTTATGCCTTTACACTCTATGAATGATTTCCAA
CCATTCTGAGGGAACCTTTGAGCGCCTCCGTTACCTTTTAGGAGGCGACCG
5   CCCCAGTCAAACTGCCCGCCTGACACTGTCTCCCACCACGATAAGTGGTGC
GGGTTAGAAAGCCAACACAGCTAGGGTAGTATCCCACCAACGCCTCCACGT
AAGCTAGCGCTCACGTTTCAAAGGCTCCTACCTATCCTGTACAAGCTGTGC
CGAATTTCAATATCAGGCTACAGTAAAGCTCCACGGGGTCTTTCCGTCCTG
TCGCGGGTAACCTGCATCTTCACAGGTACTATGATTTCACCGAGTCTCTCG
10 TTGAGACAGTGCCCAAATCGTTACGCCTTTCGTGCGGGTCGGAACTTACCC
GACAAGGAATTTCGCTACCTTAGGACCGTTATAGTTACGGCCGCCGTTTAC
TGGGGCTTCGATTCGTAGCTTCGCAGAAGCTAACCACTCCTCTTAACCTTC
CAGCACCGGGCAGGCGTCAGCCCCTATACATCACCTTACGGTTTAGCAGAG
ACCTGTGTTTTTGATAAACAGTCGCTTGGGCCTATTCACTGCGGCTCTTCT
15 GGGCGTGAACCCTAAAGAGCACCCCTTCTCCCGAAGTTACGGGGTCATTTT
GCCGAGTTCCTTAACGAGAGTTCGCTCGCTCACCTTAGAATTCTCATCTTG
ACTACCTGTGTCGGTTTGCGGTACGGGCACCTGTTATCTATCTAGAGGCTT
TTCTCGGCAGTGTGAAATCAACGACTCGAGGAAACAATTTCCTCTCCCCAT
CACAGCTCAGCCTTATGAGTGCCGGATTTGCCTAACACTCAGCCTTACTGC
20 TTGGACGTGCACTCCAACAGCACGCTTCGCCTATCCTACTGCGTCCCCCCA
TCGATTAAAACGATACTAGGTGGTACAGGAATATCAACCTGTTATCCATCG
CCTACGCCTGTCGGCCTCAGCTTAGGACCCGACTAACCCAGAGCGGACGAG
CCTTCCTCTGGAAACCTTAGTCAATCGGTGGACGGGATTCTCACCCGTCTT
TCGCTACTCACACCGGCATTCTCACTTCTAAGCGCTCCACATGTCCTTGCG
25 ATCATGCTTCGACGCCCTTAGAACGCTCTCCTACCATTGTCCAAAGGACAA
TCCACAGCTTCGGTAATATGTTTAGCCCCGGTACATTTTCGGCGCAGTGTC
ACTCGACTAGTGAGCTATTACGCACTCTTTAAATGATGGCTGCTTCTAAGC
CAACATCCTAGTTGTCTGGGCAACGCCACATCCTTTTCCACTTAACATATA
TTTTGGGACCTTAGCTGGTGGTCTGGGCTGTTTCCCTTTCGAACACGGACC
30 TTATCACCCATGTTCTGACTCCCAAGTTAAATTAATTGGCATTCGGAGTTT
GTCTGAATTCGGTAACCCGAGAGGGGCCCCTCGTCCAAACAGTGCTCTACC
TCCAATAATCATCACTTGAGGCTAGCCCTAAAGCTATTTCGGAGAGAACCA
GCTATCTCCAAGTTCGATTGGAATTTCTCCGCTACCCTCAGTTCATCCGCT
CACTTTTCAACGTAAGTCGGTTCGGTCCTCCATTCAGTGTTACCTGAACTT
35 CAACCTGACCAAGGGTAGATCACCTGGTTTCGGGTCTACGACCAAATACTC
AACGCCCTATTCAGACTCGCTTTCGCTGCGGCTCCGCATTTGCTGCTTAAC
CTTGCATCAGATCGTAACTCGCCGGTTCATTCTACAAAAGGCACGCCATCA
CCCATTAACGGGCTCTGACTACTTGTAAGCACACGGTTTCAAGTTCTCTTT
CACTCCCCTTCCGGGGTACTTTTCACCTTTCCCTCACGGTACTGGTTCACT
40 ATCGGTCACTAGAGAGTATTTAGCCTTAGGAGATGGTCCTCCCAGATTCCG
ACGGAATTTCACGTGCTCCGTCGTACTCAGGATCCACTCAAGAGAGAATAT
GTTTTCGACTACAGGATTATTACCTTCTTTGATTCATCTTTCCAGATGATT
CGTCTAACATGTTCTTTTGTAACTCCGTATAGAGTGTCCTACAACCCCAAC
AAGCAAGCTTGTTGGTTTGGGCTCTTCCCGTTTCGCTCGCCGCTACTCAGG
45 GAATCGATTTTTCTTTCTCTTCCTCCGGGTACTAAGATGTTTCAGTTCTCC
GGGTCTGCCTTCTGACATGCTATAAATTCACATATCGATAACATGACATAA
CTCATGCTGGGTTCCCCCATTCGGAAATCTCTGGATCAACGCTTACTTACA
GCTACCCAAAGCATATCGTCGTTAGTAACGTCCTTCATCGGCTTCTAGTGC
CAAGGCATCCACCGTGCGCCCTTAATAACTTAATCTATGTTTCCACCATAT
50 TTTGAATTGTTATTCAAAATAAATAGCTAAAACTAGTTATTAATCTTGTGA
GTGTTCTTTCGAACACTAGCGATTATTTATGAATTCAAGCTTATTTAAAAC
TCTATTCACTCGGTTTTGCTTGGTAAAATCTTACTTACTTATCTAGTTTTC
AATGTACAAATGAATGTTAATAAACATTCAAAACTGAATACAATATGTCAC
GTTATTCCCTCATCTTCGTAGAAGATGTTCCGAATATATCCTTAGAAAGGA
55 GATGACAAGTTCATATAATCATTGTCTAAATTTCTGATATTTGGAATTTTG
TCATGATTTAAAGTCTCAAGCAGTTTAGCTTTTTTTAATTTTTGAACAGTA
TATTCACTTGGAAAAGCTACATCATAGTGTGTGCCTCCATTTCGAATTTTC
GCTTCCATTGCTTCATTTGAATCAAAGGTTTCATATACGACTTGTATACCT
GTTTCTTTTTTAAACTTCTTAATTAAGCTAGGATCAATATATATTCTCCCCAA

TTGTAAACATATAACTTCTCATTTGAGTGGGTATTATCTTTTGATTTAAAC
CAATGACTAATCATTAAACATAAAATACCAATAACTATAGCAGAAATGATA
AGTTGAAAAAATCGCTTCATTCATTAACACCTCGCTTCATCTTTTTCTGAC
GATTAACGATATGTTGAATAAAGTAATATCCAAATACTCCAAACATAATAG
. 5   CAATAAATATTATTGTAGAAATGGCATTAATCTCCATACTTATACCTTTTC
GAGCCATAGCATAAACTTCTACAGACAATACACTAAATCCGTTTCCAGTAA
CAAAGAAGCTTACCGTAAAATCATCTAGAGAATATGTTAATGCCATAAAGA
ATCCTCCTATAATCGAGGGCATAATATTAGGTATAATAATGCTAGTTAATA
ATTGTGATTCATTAGCACCTAAATCTCTTGCTGCATTTAACATGTTATCAT
10   TCATTTCATATAATTGGGGTAAGACGATGATAACAACAATTGGAATACAGA
ACGCTATATGAGATGCTAGCACTGTTGAAAATCCTAAACCTAAACCTGTGA
AATGTCCTATTGCAGTGAACATAATTAAAAAAGAAGCACCAATGACTACGT
CTGAAGATACCATTAATACATTATTCATCGTTAATAGCGTAACCTTAAACC
GTTTGTTGCGTAAGTAATATAAAGCGATCGCACCAAATGTACCTATAACTG
15   TAGCAATCGATGCAGATAAAAGTGCCACTGCAACTGTATTAAAAATCACCG
ACATCAACCGATCATTATGAAAAAGGGAGTGGTAATGTTCTAAAGTAAAAT
GTTCGAAGTGACTCATATTACCCGCAGAATTAAAAGAATAAATCATTAAAA
AGAAAATGGGAATATAAAGCACACTTATTAATAGAGCTATATATAATTTTC
CGTACCATTTCACACTAATCACCCCTTCCCATTAGATGATTTTGATTTTGT
20   AATAATGAGTAAAAAGGCCATAAAAATAATGAGAAAAGTGCTATAGTGGA
ACCTAATCCATAATTTTG

Sequence 3459
25   step.1003e03.cons.ok
TTTTTGTAATTTTACTGCTTCAATCATTTTCACAATAAGTTGTTTAAATTC
ATTGATATTAAGGTGTAATTGAGTGTCATCTTTAGCGTGGTTAATAGTTAA
TAAAGCTTTTTCAAAAGATTTTAATGCGCCTTTTAAATTCCCACGACGATA
GTGATAATTAGCTGTTGCAAACAAAATGAGACTTACTACTGCATCATGTTT
30   TGAAAAAGTATTTTGTGATTTCCAAGCATCTTCTAAAATATCATGACATAG
GAAGTAATGTTGATGCTTATGAAATTGATAATAATAATCCATTAATGATTG
ATCCATTAGATTTCACCCATAATTTAAAAGTTGTATGTATTCATGCTATAA
TAATTTAGTGAAATATGCATTAACGATGCAATTTGATATAGAGGTAGATGT
AATGTATGAAGTAAAACTCGATGCATTTAATGGTCCATTAGACTTATTATT
35   GCATCTAATTCAAAAATATGAAATTGATATTTATGATATCCCTATGAAAGC
CTTAACTGAACAGTACATGCAATATGTTCATGCGATGAATCAGCTAGAAAT
TAATGTTGCTAGTGAATATTTAGTTATGGCATCAGAATTACTAATGATTAA
AAGTAAATTCAACGATGAATATAACCAAAAATCCAACACTAAAAACATGCT
GTGCGGCTTCTGCACTATGAACTGTTGTTTCACCATTAACTGTTGTAGTAA
40   AATTAGCATCCATCGCTCCAGTAATTGAAAATACCAAGAAAACTAAAAACA
ATAGACTAGAAAAACTTTGTAAAATGATTCCTATCAATGATGTTATATTAA
TTGCTTTATTCATCATAATTCTCCTTTTCATTTTCTACGTATATCCTTACAA
TCAGGTTTATGTTCTTTAATATAATCTTTAGGATTTATATCTCGTTTATAG
TAGACTTTCTTTCCCTTTATTTCATATGCAACATAAAATAGTTGCTCTTTA
45   TTTTTGAATAATTTATAACTAATAATAATCAATTTTCCATATAAAGAAGTA
GATTTTTCCGCTTCACCAGAAATAATTAAAGGTGTTCTTGCTTCGTCGATC
AATATGGAATCGACCTCATCAATAATTGCAAAATGTAGAGGACGCATTACT
CTCTCTTCAGCATAGTTCACCATATTATCTCTAAGATAATCAAACCCAAGT
TCATTATTCGTACTATAAGTGATATCTTGTGCGTAAGCCTCACGTTTTTCT
50   TCAGTTGACTTACTATTTAAGTTCAAACCTACAGTTAAGCCAAGATAGTTA
TATAGTTCAGCCATTTCTTCACTTTGTGAACTTGATAGATATTCATTGACT
GTAATAACATGTACACCTCTACCAGCTAAAGCATTCAAATACGTCGGCATG
GTTGCAGTCAATGTTTTCCCTTCACCTGTTCTCATTTCTGCAATATCACCT
TTATGTATAGCAATACCACCCATTACTTGTACTTTATAAGGAATCATATTA
55   AATACTCTCTTTGACCCCTCACGTACAAGTGCATAAGCTTCAGGTAAAATT
TTATCTAAATAATCATTTTGTTTTTTTTACATCTTGATGATTTAGTGATCAC
CACAATCCTTTTATCATGTAACTAACCTAGTAAATTTAAGTTTCAACTTTT
ACTTATAGTGTTTGACTTTAATTATAAATCAAAATTCGTCATAATAAGACG
TGCTAAATTTACAATAGCTAAAACATGATAGAGGTGACTTTTATGTCTACG

```
AATAATGAGATTGAATTTAAACAAATACTAGATCAAGATACTTACTCAAAA
ATCTATGAACACTATTTCAAAAATCAATCACCTTTTAAGCAAACTAATTTC
TATATCGACACAGAGAATTTTAAATTAAAACAGCATCATGCTGCTTTGCGT
ATAAGGGTAAAAGATTATATGTTTGAAATGACTTTAAAAGTCCCTGCTGAA
GTTGGATTGACAGAATATAATCACTCAGTAAATATAGAACCTGAACTTGAT
ATGTCACTTCAACTTTCTCAATTACCCAACGATATTAGAAATATTTTAGAA
CAGGACTTTAATATTTTAGAAAATGAGCTTAAAGTACTAGGAAACTTAACT
ACCTATCGTTTAGAAACCGATTATCAAAATGAATTACTAGTATTAGATAAG
AGTGAATATCTCGGCAAAACTGATTATGAATTAGAGTTTGAAGTTCATTCT
TATGATGAAGGATATTCAAAATTTAAAACTTTACTTCAACATTTTAATCTT
CAACATCAAAAACCCTTGAATAAAGTGCAACGTTTTTTTCAAGAAAAACAA
AATGCAAGTGATAAAGAGTAACAGTTGTCGATAAGGCAAATTACTTTTAAA
TGTAAAAATTCATGTTATATTAGATATATATTAATGGAACACGGTGATATT
ATGTCTAAAACACCATATGAGTTGATTGGTCAAAAAGCCTTGTATCAAATG
ATTGATCATTTCTATCAACTTGTCGAGAAAGATTCTCGTATCAATCATTTA
TTCCCAGGCGATTTCAAGGAAACCAGTCGAAAGCAGAAGCAATTTTTGACA
CAGTTTCTTGGAGGTCCTGACTTATATACCCAAGAACATGGTCATCCCATG
TTAAAACGAAGACATATGGAATTTACAATTAGCGAGTATGAACGTGATGCA
TGGCTTGAGAACATGCATACTGCTATTCAACACGCCAAACTTCCTGCGGGT
GTAGGCGATTACTTGTTTGAGCGATTAAGACTTACTGCAAATCACATGGTA
AATTCCTAAATTTAATTGTAGGTGAAAATACATGGCTGAAGAATTAAGAAT
CATGGAGAATAAGAGTCGTGAAGATACTAATCTATCACCTGTTAGCAAAAT
AGAAATCTATTCTTTTTTTGATCCTTTTAGCAAAGATTGTTTTAAATTATC
TGCAATCTTATCAAAATTAAGAATTGAATATAATAAATATATAAAGGTAAG
ACATATTTTAAACCCTTCTTTAAAGGTATTAACTAAGTGTCAAGCTCAAAG
TACTTCAGATTTTGACAATATTGCACTTGCCTATAAAGCCGCTGAACTTCA
AGGTCGTATCAGAGCAGAAAGATTTATACATTTAATGCAAAATGAAATCAT
TCCAAAACGTGATATTATTACCGAAGATATGATTTCTGATTGTATTAATAA
TGCCGGCATTGACTATCAAGTTTTTAAAGAAGACTTGCAAAAGGACAAGTT
GACTGACAGCTTGAAAGTTGATCTTCACATTGC
```

```
Sequence 3460
step.1003e04.cons.ok
TATATCGGCACTGAAAAGCCTATGCTAAAAAATAAAAATATCGCGCTTCTT
TTTGAAAAAGATTCCACTAGAACACGTTGCGCATTCGAAGTTGCCGCACAT
GATCAAGGTGCACACGTCACTTATCTTGGACCTACAGGTTCTCAAATGGGT
AAAAAAGAAACTGCTAAAGATACAGCACGTGTACTTGGTGGTATGTATGAT
GGTATTGAGTACCGAGGTTTCTCTCAACGTACTGTAGAAACATTAGCGCAA
TATTCAGGTGTTCCGGTATGGAATGGATTAACCGATGAAGATCACCCTACA
CAAGTGCTTGCTGACTTTTTAACTGCTAAAGAAGTATTGAAAAAAGAGTAT
GCTGATATCAACTTTACTTATGTTGGCGATGGACGTAACAATGTTGCTAAC
GCATTAATGCAAGGTGCTGCCATTATGGGTATGAATTTCCATCTTGTTTGT
CCTAAAGAACTCAATCCGACAGAAGAATTATTAAATCGTTGCGAACGTATT
GCGACGGAAAATGGCGGTAACATTTTAATAACAGATGATATTGATAAAGGC
GTGAAAGATTCTGATGTTATTTATACAGATGTTTGGGTATCAATGGGCGAA
CCTGATGAAGTATGGCAAGAACGCCTTAAACTTTTAAAACCATATCAAGTT
AACCAAGCATTATTAGAAAAAACAGGCAATCCAAATGTTATTTTTGAACAT
TGTTTACCTTCTTTCCACAATGCAGAAACTAAAATTGGTCAACAAATTTAT
GAAAAATATGGCATTAGTGAAATGGAAGTCACTGATGATGTCTTCGAAAGC
AAAGCTTCTGTAGTATTCCAAGAAGCTGAGAATAGAATGCATACAATTAAA
GCGGTCATGGTAGCAACTTTAGGAGAATTCTAAATTATAGAAGGAAGTGAG
TGAAATGGCTAAAATTGTAGTAGCTTTAGGTGGAAACGCTTTAGGAAAATC
ACCACAAGAACAACTTGAATTAGTAAAAAATACAGCTAAATCCCTAGTAGG
ATTAATTACTAAAGGTCACGAAATTGTGATTAGTCACGGTAATGGACCACA
AGTAGGAAGTATTAACCTTGGTCTGAATTATGCAGCTGAACACGATCAAGG
TCCTGCTTTTCCATTTGCTGAATGTGGCGCTATGAGTCAAGCCTACATCGG
CTATCAACTTCAAGAAAGTTTACAAAATGAACTTCATTCAATGGGCATAGA
TAAGCAAGTTGTCACACTAGTTACCCAAGTAGAAGTTGATGAAGGCGATCC
```

AGCTTTTAATAGTCCAAGTAAACCCATCGGTCTGTTCTACACTAAAGAAGA
AGCAAATCGTATTCAACAGGAAAAAGGTTATCAATTTGTAGAAGATGCTGG
TCGAGGTTACCGTCGCGTTGTACCATCACCACAACCAATATCTATTATCGA
ACTGGAAAGTATTAAAACTCTAGTAGAAAATGACACACTCGTCATCGCTGC
AGGTGGAGGTGGTATACCAGTCATTCGCGAACAGCATGATAGCTTTAAAGG
TATAGATGCCGTCATCGATAAAGACAAAACAAGTGCATTATTAGGTGCTGA
TATTCACTGTGATCAACTCATTATTTTAACAGCGATTGATTATGTTTATAT
CAACTATCATACTGACCAACAACAAGCACTTAAAACAACAAATATAGATAC
GCTTAAAACATATATTGAAGAAGAACAATTTGCCAAAGGCAGCATGCTACC
TAAAATCGAATCTGCCATCTCCTTTATTGAAAATAATCCTAACGGTAGCGT
GCTCATCACATCATTAAATCAATTAGATGCAGCACTAGAAGGTAAAATTGG
CACACTCATTACAAAGTAACGTTTTAACGTATAAGTAGAAATTAATTTAAT
CATATTCATACAGTGCATCCTGTTATTCCACTCCTCTCCATTTCATACAGG
GTGCACTTCCTTAATCAATCATGTAAAATACTAACCAAATTCAATTGATAG
GCTTTTATTTTAGGTTAACTGTCGAAGGAGATGAAACCGTTGGAACAAGCG
ATCAATGATAATAAAAAGAAAAAACGTTTTAACTTTAGAATGCCAGGTGCA
TTTATGATTCTCTTTATCCTAACAGTTGTCGCAGTTATAGCAACTTGGATA
ATCCCCGCGGGTGCATACTCAAAACTTTCATATGAACCTTCATCCCAAGAA
TTAAAAATTGTCAATCCTCATCATCAAGTAAAAAAAGTTCCTGGAACACAA
AAGGAGCTTGATCGATTAGGAGTTAAAATCAAAATAGAACAATTTAAATCT
GGTGCAATTAATAAACCCGTTTCAATTCCTAATACTTACGAACGTCTAAAA
CAACATCCAGCTGGTCTTGATCAAATTACTAGTAGCATGGTTAAAGGAACC
ATCGAAGCCGTCGATATTATGGTCTTTATACTTGTTCTAGGTGGACTGATT
GGTATTGTTCAAGCGAGCGGGTCCTTTGAATCAGGATTGTTAGCACTTACT
CAAAAAACGAAAGGCCACGAATTTATGTTGATTATGTTCGTAGCAATTTTA
ATGATTCTGGGTGGAACACTATGTGGCATTGAAGAGGAAGCTGTAGCGTTC
TATCCTGTACTCGTTCCAATATTTATTGCGCTTGGATATGATTCTATTGTC
TCAGTCGGTGCAATTTTCTTAGCAAGCTCTGTGGGTAGTACATTCTCAACA
ATCAACCCATTCTCAGTCGTCATTGCTTCTAATGCAGCAGGAACAACTTTT
ACTGATGGTCTTTATTGGAGAATAGGCGCTTGTATCATCGGTGCCATATTT
GTTATTAGTTATTTATTCTGGTATTGTAAAAAAATTAAAAAAGATCCTAAA
TCCTCTTATTCTTATGAAGACAAAGCAGCATTTGAAAAACAGTGGTCTGTG
CTCCATGATGACGGTTCTTCTGAGTTTACATTACGTAAAAAGATTATTCTT
ACGCTTTTCGTCCTACCATTCCCTATTATGGTTTGGGGCGTCATGACACAA
GGATGGTGGTTCCCAGTCATGGCATCTGCATTCTTGATCTTTACCATTGTC
ATCATGTTTATTGCTGGAACAGGACAATATGGTTTAGGCGAAAAAGGCACT
GTAGATGCATTCGTTAATGGCGCTTCAAGTTTAGTAGGTGTATCTTTAATC
ATTGGTTTAGCTCGAGGAATCAACTTAGTATTGAATAAAGGAATGATTTCT
GACACAATCTTGCACTTTTCATCATCTATCGTGCAACATATGAGTGGGCCT
TTATTTATCATTGTTCTGCTCTTTATCTTTTTTCTGTTTAGGATTTATCGTG
CCGTCCTCATCAGGATTAGCAGTACTATCTATGCCTATCTTTGCGCCATTA
GCTGATACAGTAGGTA

45   Sequence 3461
     step.1003e05.cons.ok
     CTATTTCATGATGATCTGTCACTATAACATCTACACCTAATTCTTGTATGG
     TACTTATTTCATGATGTCCCTGTATACCATTATCTACAGTTATTATTAAGG
     AAATCCCTTCGTCATAAGCATTTTTAAATGCTAATTCATTAGGTCCATATC
50   CTTCTGTAAATCTATTGGGAATATACCACCCCACCTGAGCGCCAAGTAGAC
     GTAAAGTAGACACTAGAATTGTAGTAGATGTAACACCATCTGCATCATAGT
     CACCATATACTAGTATTCGTTCATTTTGATCGATAGCTAATTAATACGAT
     CAATGGCTTTTTGCATATCACTTAATAACATATAATCATGACTAATGTTAG
     TGCCTTCAAATAGATTTTGAATCTCCTGATTATCCACAATATTTTTACTCT
55   CTAAATACCTTTTTTTACTATAGGTGAAAGTTTAAATTGATCACTGCTACTCT
     CACTTATAAAGGTATTCGGTTGAGTTATATCCCAATTATATTTTGATTTAA
     TCATTATAATCGTCCTCATTTCCACAGGTCATTATACCGAAATTTTATTGA
     AATAAAAAGTACTTGCATTTTTTAAAAGAATATATCATTTCAATTCTTCAC
     AAAACTTTTATCATTCAATTTGTTCATTAACTGCTATCTTTATGATTAAAC

GCTAATATATAAAATCAAGAGGCAGGGGCATAATTCCTAGCAAAATAGACA
GTAAATGAGTTTTTATAAATTCATTTACTGGCTTCTTTATTTACAATACCT
CGTATTGTTGGCTCGCTTTCTTAGGGGACAGCTTCAGCCTGTAGTCTTCAA
CTTGTCCTGCTCCCTCAATAGTCTCGCCAAAATACTTTGTGTTTATATGTA
5   ATTTCACATTGAAATACTTTAAAAAATAAGGTCTTTCGTATAATTTAATAA
ATAGCAATAAACTAAATTAACGAGGTGCCTTATGTATAAAGATTATAATAT
GACTTAACTTACTCTACCAATGGAAATTTCAGTTCTTATCCCCACACATGA
TATTTCACGACATGTAAATGATATTGTAGAAATAATTCCAGATACAGAATT
CGATGAATTCAGACATCATCGTGGCTTAATATAAAAAATAATAGCTCAACG
10   AGCTGAAAATAATCAAAAAATTATAAAAAAGACAATTTCTATATTATTTCA
ATAGAAATTGTCTTTAATTACTTATCTTGGACCTTTTTGTTCCAGCCTCTT
GTAAATTTTTAAAATTATTAAAATGTGAAGAAGTTAAACTAAGATTTTTTC
ATCGTTAGATTTCTTCTCTTTGTGTACTACTAATTTATTATTTTTTGTCTT
TTTAAACTGTCGTTTCTTAAGCATGCCCCATAATGGTACAGCAATGAAAAT
15   TGACGAGAATACACCAGATAATAATCCAATTAGTAATGCTAAAGAGAAATT
AAATATTGTTGGTGCACCTAATATTAATATTGCAACTACAACTACAACTAC
AGTCAACACTGTATTAATAGAACGTGTCATAGTTTGTCTAATAGAGCGGTT
GACTATATCATCAATTTGATCAGTATGCGTAATTACTTTAACTTTATGCAG
ATTTTCTCGAACACGGTCGAAAGTTACGATTGTATCATTGATTGAATAACC
20   AACGATAGTTAATACTGCTGCAATAAATGTTAAATCTACTTCTAATCTAAA
CAAACTAAAGATTGCAATGATGATAAATACGTCATGTAATAATGCAAGTAC
AGAAGATAGACCCATACGCCATTCAAATCTTAGTGAAACATAGATAATAAT
GCCTATAGAAGCAAGTATTAATGCAGTTACAGCATTTTTAGCTAGCTCTTG
GCCTATGAGAGGTGAAACGGTATTAATTTGTGGATTATCTCCAAATTCAGA
25   TTTCACCTTAGCACTTAATTTATTGTCTTCCTCACGTGATAAATCTTTTTT
AAACTGAACTGTGGCATTTTTATTTCCACTACCATTAATCTGTATTTGATC
TGGTTCCAATCCAACTGATTTTACAGTTTTCTCAACCTGTGCTTGTGTTAT
AGCATTTTTAGATTGAATATCTGCTCTTGTTCCGGATGAGAAATCAATACC
TAAGTTTAATTTAAATATTGAAATGATAATCAAACCAATAATTACAATTAA
30   AATACTAAGTGAAATAAGTGGCTTAGCTAATTTAACAAAGTTTAACCTTTC
ATATGATGTTTTTAAATCATGTACATCTTTACCTTCATTAATATCATGTCT
ATCCTTCTTCTTAACACCAAATAACCAGTATTGTTTTTTAAAGAAGTTTGA
AGATACCAGTAATGATAACAACCCTCTTGATAAGAATACTGCGGTTACAAA
TATCATTAAAATACCTAAGAGTAACATGGTTGCGAAGCCTTTGACTGAACT
35   TTCTCCAAAGAAGAAAAGCACAGCTGCAGCGATGACAGTTGTTAAGTTGGA
ATCAAATATAGTTAAGAATGAACTTTTTATTTGCTTTTGAATACGCTTGTTT
AAGCGTGCGTCCAATTCTTAGTTCATCTTTAATACGTTCATACATTATGAT
ATTGGCATCGACAGCCATACCTACACCTAAAACTAATGCCGCCAATCCAGG
TAGAGTTAATACACCTGATATGAAATTGAATGCGACTAAAGTTAAATAAAT
40   ATAAGTGGTTAAGGCAATGATTGCAACTAAACCAGGCAAACGATAGAAACC
AAGCATAAATAAATAAATTAATGCTATACCTACAATTGATGCAAACATGGT
CTTATCAAGAGCATCTTGACCAAATTGTGCACCAACAGAGTTAGAGTAAAT
TTCTTTTAAATCAACTGGTAATGAGCCGGCATTTAATAACTCAGCTATTTG
TTTCGCTTCTTCAACACCTTTTTTCCCATTGAAGCCACCTGAAATTTCAAC
45   ACTACTAGAATTAATAGGTTGGTCTACACTCGCTGCAGATATAAATTTAGG
CTTTTTACCTTCTTGTTGCTTTTTAGCTTCTTTCTTGTAACTATCGCCTTT
TTCGAAATCTAACCAAACTACCCATGACATTGTCACGTTTTTTAGAAATCTT
TTCAGTTACTTTCTTAAATTTATCTTTACTTTTTACTTTAAATGTAACTGT
TGGTTGATTAGTTTCTTGTTTAAATTCTTGTTTAGCAGAGCCTTGTTTAAT
50   GTCTGAACCAGACATTAAAACATGATCTTCAGCATCTCTAATTGTTAAATT
AGCTTGTGTCGATAATAATTTACGCGCTTGTGCTTGATCCTTGATACCTGC
TAATTGTACACGAATTCGATTTGGATCTTCGATTTGTATTTTCGGTTCTGA
TACACCTAGAACATTTACACGATTTTCTAATGTTTGAGATGTAGCTTGAAG
TGCTTTTTTATCAATTTTATCTCCTTTATTTAAAGGATCTACTTGGAAGAG
55   GACTTCAAAACCACCTTGCAAATCTAGACCTAAATTAACATTTTTAACTAC
ATTCTTATACGTCATACCCTTTAGAGCACAGTGGCGATGATAT

Sequence 3462

step.1003e06.cons.ok
```
ATTTCTTAAAATGTTGTAAGTCTTATGTTTGCTTTTATTTCACGAATCTTT
TTATTTATTTCAAACAATTGTTTACATGTGATTTTAAATCATCTCTTCTAA
ATACAATTAAGAATACTATACCTGTATCTGACATTCGTCAATAGTTAAATG
AAACTCTTTTCACATCGTCATGTATCATTCATTATTTTAGTAGACTGATAG
CATAATGTAGCAGTTTTATTTATTTAATTAAAAATATTAAATGGAAATAGA
ACGTAAGGTGACATTACAAGCAGTTAATTATAATAATGTTGAAATTCTCTC
GCCTCTCATCCGAGAAAAATTGTATGTATATCGTCACAATGTATCTACTAT
TACACAGAACAAAATATTTAGTTTCAACAAGTTGTAGAAGCGGGACACTAA
TTATTTAAAAAGTGTTGCATATAAGCACATCTAAATATGACGTAATTGTTG
CAAATTTAAAATAACTTGAAATGTGATAATGACTTTGCTATATTAAAATTA
AGTTAATAAAGTATTGTTCGTAGGACAAGTAACATAGATAGTTCGATTTCA
GAGAGCTTGTGGTAAGTGAGAACAAGTAATCGACATTCATGTGAATCTACC
TACTATATGTGAACAATCGGTAATAACCGTTATTTTAGTTAAGCGCAATTT
GAGGTAAATGTTTTTAATTTACTTGAAATTGTTAAATAGGGTGGCAACGCG
TAGACCACGTCCCTTGTCTGGGATGTGGTCTTTTTTTATTGTTTTATCACA
CGAAGTCATCCATAAAATTGAATATATTTTATTTGGGAAAGGATGAAGGTT
ACATGTTAGACATTCGTTTATTTAGAAATGAACCTGAGAAAGTGAAGAGCA
AAATTGAATTAAGAGGCGACGATCCTAAAGTTGTCGACCAAGTTTTAGAAT
TAGATGAACAACGCCGTGAATTAATCAGTAAAACTGAAGAGATGAAGGCGA
AAAGAAATAAAGTGAGCGAAGAAATAGCTCAAAAGAAACGTAATAAAGAAG
ACGCTGATGATGTCATTGCTGAGATGCGTCATTTAGGTGATGAAATTAAAG
ATATCGATAATCAACTTAATGAAGTAGATAATAAAATTAGAGATATCTTAA
TTCGTATTCCTAACTTAATTAATGAAGATGTACCTCAAGGTGATTCTGATG
AAGAAAACGTTGAAGTTAAAAAATGGGGTACGCCACGTGATTTTGAATTTG
AACCTAAAGCGCACTGGGATTTAGTTGAAGAATTAAAAATGGCTGACTTTG
AACGTGCTGCTAAAGTATCTGGTGCTCGTTTCGTATACTTAACTAAAGATG
GCGCATTACTTGAACGTGCTTTAATGAATTACATGTTGACAAAACATACAA
CGCAACATGGTTATACTGAAATGATGACACCTCAATTAGTGAATGCTGATA
CGATGTTTGGAACAGGTCAATTACCTAAATTTGAAGAAGATTTATTTAAAG
TTGAAAAAGAAGGCTTATATACGATTCCAACTGCAGAAGTACCTTTAACAA
ACTTCTATAGAGATGAAATTATTCAACCAGGTGTACTACCTGAATTATTTA
CAGCTCAAACTGCATGTTTCCGTAGTGAAGCAGGATCAGCTGGTAGAGATA
CTAGAGGGTTAATTCGTTTACATCAATTTGATAAAGTTGAAATGGTTCGTA
TTGTACAACCTGAAGATTCTTGGGATGCTTTAGAAGAAATGACACAAAATG
CTGAAGCTATTCTTGAAGAATTAGGTTTACCATACCGTCGTGTTATCTTAT
GTACTGGCGATATTGGTTTCAGTGCTAGTAAAACATATGATTTAGAAGTTT
GGTTACCAAGTTACAATGATTATAAAGAAATCAGTTCTTGCTCTAACTGTA
CTGATTTCCAAGCACGTCGCGCAAATATCAGATTCAAACGTGATGCTGCTT
CTAAACCAGAATTAGTACACACATTAAATGGTAGTGGTTTAGCAGTAGGTC
GTACATTTGCAGCCATCGTTGAAAACTATCAAAACGAAGATGGTACATTAA
CAATTCCTGAAGCATTAGTACCATTTATGGGTGGCAAAACTAAAATTGAAA
AACCAATCAAATAATCATATAAACTTAACTAGATGATTACAATATTAAAAA
GTTATTTAACTGTCGATAAGATTCTATAAACTTATCGACAGTTTTTTAATA
TATAAAGGCGAAATTGGTATTTTTATATTCAATTAGGGATGGATGTCTAAA
ATAGTTTAAAATGATTTTCACGATTTAAAATACTGGTATAAAAGCGATTTT
ATAGATACGGTTGTCTGACGTCACACAAACGAATGTCTAAAAAATAGTGAA
TATTTATTGTCATAGATTGAGTTATAAAATCATGAAAATAGTACATGAAGT
AGCAAGTATTTATCTTTTAGATAACGCCTGTAAAAAATATAAAATAGTGGAA
ATTGATTTTTTAAACGTTATCATTCATATTCAAAATATTTAACTTAGTTTG
TCAAACAACGTTTATTTTTATATAACCATAAAATAATATCATAAAGAAGCT
ATATGAAATTCGAAGTATAATGCCAATGTTTCATATAGCTCCTATTCTAAA
TATTAAAATTTAATAATTAAATGTATCAATCGTCCTATTCATTTGATGCAT
AGCCTGAGAGTGGGTAATGATCTGAGTAATCATTGTATTTGTATGTTTTTC
CCCAAGATTTCACGGACCATTCTGGTGACTTCACTCTATGTGTATCATTAT
GCCATGAGCTTGGTTGTGCATGGTCATGATCTAATAAAATATAATCTAAGT
GTTGAGGTTCTAATTTAGGATAATTATATTTCGCAATACTATTACTGCTAG
TATCCCAACTATATGCATTACCATCGAATTGAGTAGGTAATGAAACATTTA
AGTTATTTGCCATTTGTTGATATTCATCTGAATCTTTAATGACATTTAAGT
```

CACCACCGATATAGACGGGTTCATTTTTAGGGATATTCTTGTCTTTGATAA
ACTGTTTAATTTCACTCATTTGACTCTGTCTAATGTCTTTATCTTTTCCTT
TAAAGCATGTTGGATCTTCAGCTTGTAGATGTGTTCCGATAAT

5

Sequence 3463
step.1003e08.cons.ok
ATGCGTTTATCAAACGTCTCATTAAACAATTTGGTAAACCTCAAAAGGTAA
TTACAGATCAGGCACCTTCAACGAAGGTAGCAATGGCTAAAGTAATTAAAG

10   CTTTTAAACTTAAACCTGACTGTTATTGTACATCGAAATATCTGAATAACC
TCATTGAGCAAGATCACTGTCATATTAAAATAAGAAAGACAAGGTATCAAA
ATATCAATACAGCAAAGAATACTTTAAAAGGTATTGAATGTATTTACGCTC
TATATAAAAAGAACCGCAGGTCTCTTCAGATCTACGGATTTTCGCCATGCC
ACGAAATTAGCATCATGCTAGCAAGTTAAGCGAACACTGACATGATAAATT

15   AGTGGTTAGCTATATTTTTTTACTTTGCAACAGAACCTATAAAAAACATGG
GCTTAAAGGGTTAATTTATATAGTATTACTGCAAACATTGATTTGGGTCAA
AAATTAGGGGTATTAAAAAAATGAATGAATCATAATTTTATCAAGCTGATT
GGAGAGGTTAAAATGCATTATATAAAACTTTAGAGTATCATTAATATCTTC
ACCTCAGATATACAAAAAACTACTAGCATATAAAAGATATATCTTTTATAT

20   GCTAGTAGTTTAAAAGCTTAAACAATCAATCGTCCATTTTATTAATATATT
CTTTTTCTTTATAGTATAATTTTTCTCGTCTTTCTTCTCGTAGTTTTGCAC
GTTTTTTTAATTCTTCTGGCGTATTAAACTCTTTTTGGCTTAATAAGATAG
TAGATTGTCTAGATGCATCGTCTTTCCTGTACATATATGCACCATTACGGT
AGGCCGCTCTACTTATCAAATGCATTCCAACCGGAGATGTAAGATTGATAA

25   AAACTAGTGATAATAATAATCTGACACTGAAAAAACCTGAATTCACAATAA
AATAGATCAGTACACCAACTACAGTTAGTAATACTGACAATGTAGAACTTT
TCGTTGAGGCGTGACTTCTTAGAAAGACATCTTGAAATTTTACAATCCCTA
TTGCACTAATTAATGCAATAATACTTCCTAAGAAAATAAGTATCGAAGCAA
TAAGACTAACGATGTCTTTTATGATTTCCATTGAAGACACGTCCTTCCCCG

30   ATGAATCTTGAAATTGAGACCGAACTGACAAACGAAATAATTGCAATTAAC
ATGATTGAATCTAAATAAGACACTGAGTTAAAAATAACGCTCATCACACCA
ACAATAGACATAACAACAGCACTCGAGGCATCAAATGATACAACTCTATCA
GCAGTAGTGGGACCTTTAATTAATCTGACTAGACAAACAAGTAGTGCCATA
CCAAAAATCACTAATGCACTTATAATAAATATTTGAGTGAACATTTCAATC

35   ATCGTGTCACCTCCAAAATTAAATCCTCATACTGCTTAATACTTTTTAGAA
GATTTTCTTTATCTTTTTCTGACACATCAATACTGTGAATAAAAAATTTAT
TAGTATTTTTAGAAATTCGAATAACTGTCGATCCTGGAGTAATAATAATTA
AAATCGTTAAAAAAGTAATAGCCCAATTACTTTTTAATGAAGTTTCATATG
TGAGTAAACCTGGATTAACTTCATTCGTCTTAAATAAGATGTAATTTATGG

40   TACTTATACTAGAAGTAATAAGCTGGTATAGGTATACAGCTAAAAATTTAA
TAGCCACCCATATCTTTTTCAAATAAAATTCTTCACCAAAAAAGCGATGCA
GAATATAAATCACAATTAAACCAATTAAAAATCCAGCAAAGAAGGTTGTAA
ATTTAAATTCATCTTCATCTTGAAAGAGTACCCAAAGGAACGCGATAACAA
TATTTAATACAACTTGCTTCACTTATGATTAACCTCCTTCAAATGTGTATT

45   TACATTCTTTTGAAAGACATCTTCTTTCATATTAAGATTTGTTGCATCCTC
TGTTACTTTCAGAACAACAGGGGCTGCAATACCCATCGCTAATACCACTAC
AACTAAAACACTAAGTAAACCTTTACGATAAATAGGTAGTGGTCTAAATTG
TACTTGTTCTCCGTCAGCATCACCAAAATACATTATAAACATCACTCT

50

Sequence 3464
step.1003e09.cons.ok
GATATTGAAAAATAGGCAATAAAGTTATTAAAAACATAATAACTAACGTAG
CACTATGTTTAGACATTATTTTTAAAACCAACACCCCTCCTTTCATATAAA

55   AATTTTGCTCCGTTTTTTCATATTAGACATTATCACTCATCTTAACAATAAC
TATTCTAATTTTTTTCGAATTAGACACAAATATAAAAATAGAAATTTTATCT
AGTAAATTAAATACAAGTATATATAAATCATAAAGTTAATAAATCAATTAA
GTATGTAGCAAAATGGCAATTTGTAATATAAAATAAAGTTTATTCATTTTA
GAATCACCTTACTAAGTGATCAACAAAATTTCGAAGATTGAATAAGTTGTA

986

```
     TAAAGTTTATTATAATTAACCATACTATTAAATCATAAGTAATCTGCATAT
     GCCTTAAAGATAAATTATATATTTAATGATTAATGTACAATCTAAAGTTAT
     TAGTAGCTGGATTGATAATATAGCAAAATTCATCTACTTTTTCTTATACCT
     AACTATACAATAGGAATGTGATTCAACTTTATATATCATATTTAGGCCAAG
  5  AGAGTATGATATTTAAAAATTATGTCCGAAGTTTAAGAAGCTTGATGCCAC
     AAATTTAGAAGTTTAAAATTTATTTAATTAACTTTAGTATCATCCTTATTC
     TAGTGGACGCTGAAAGGAACTATTAATTATGATTACAGCTTATAAACACGC
     ACTTCATCATCCTTCACAAATTATTGAAACTGAAATCAATCATAGTGCTTC
     ATGGATAAATGTAGTAGAACCAGATAGAGAAGAAATTGAAGGTTTGATGGA
 10  ATTTTACAATATACCTGAGGATTTTATTCGAGATCCGTTAGACTCTGAAGA
     GAGTGCACGTATTGAATATGACGAAGATACGGGATACTCGCTAATTATTAT
     TGACTTACCTATCGTCAATTCTACCAATCGGCGTGTCTTGTCGTTTGTGAC
     CATCCCATTAGGCATTATTATTGGAAATGGTATCGTTATGACTGTTTGCGA
     TGCTGAAAATGAATTTTTAGAAAACTTTGCTAAACAAGAAGATATTAACTT
 15  GAAATTTCACAGTCGATTTGCACTTGAAATACTAACAACAATAGCAAATCA
     CTACAATAGAAATTTAAGATTGCTTAATAAATCTAGGATTCGTATTGAAAG
     AGAACTCAAAAATAACATTACTAACAAGCAGCTTTTTAAACTAATGGAAGT
     AGAAAAAAGTTTAGTATACTTCTTAGCTGCACTAAAAGGTAACGACACAAT
     TATTAAAAAACTCTTTCGTCTTCCTGCAATCAAACGTTTTGAAGAGGATGA
 20  AGAGCTACTTGAAGACTTGGTGATAGAAAACAACCAAGCTATTGAAACGAC
     TGAATTACATCAACGAATACTTGAGAGTATCACTTCTTCGTATGCTTCATT
     ATTATCTAATGATATGAATAACATCATGAAAACATTAACGTTGTTTACGGT
     TCTTTTGACCCTTCCTACACTCGTCTTTAGTTTCTTTGGTATGAATGTACC
     CTTACCAATTGATGATCATAGTTACGTGTCTTGGATTATTGTTGTGGGAAT
 25  TTCACTTATTCTAGTAGCTATCGTTAGTATCTTTTTATGGAAAAAACAAAA
     ATTATAACAACAAATACCCTTAGTTATAGATTAGCACTCAATCTAACTAAG
     GGTATTATTATATGTTTAAAATTGTGTCAAAATTTTAAGTCTCTACTATGA
     TTCAACAGCTGTCTTATTTGTTTTTCTTCGAGAAAGTACTAAAGTGACGAC
     AACTGAAACAACAAATGCAATTAGCATTCCTATTAAGTAATGTAACCAACC
 30  AGCGTGTGTAGGATTTATAGATATAAATCCAGGTAATCCAGCTGTACCAAG
     CGCTATCGCTTTTACTTTGAAGAATGAAATATAAGCCGCACCTATACCTGA
     TCCTGCAACAGCACCTATAAATGGATATCTCAATTTAAGATTGACACCAAA
     CATTGCTGGTTCTGTAATTCCTAGTAAAGCTGAAATACCCGCCGCGGAAGC
     AACACCTTTTAATTTTTTATTTTGCTTAATGATAAAGAATGCAGCTAAAGC
 35  TGCACCACCTTGTGCAATATTTGACATCGTTGCGATTGGGAAGATAAATGA
     ACCACCTGTTTTAGTCGCATCAGCAATTAATGTCGTTTCAACTGCAATAAA
     GCTATGGTGCATTCCTGTAATGACGATTGGCGCATATAATAAACCAAATAT
     AAGTCCACCAATAGCTCCACCAAATTCATATAACCAAGTCAATCCATCAGA
     TAACCAATAACCTAATTGACGAGTGACAGGCCCTACAAATAAAAATGTTAT
 40  AAATGCTGTAATAAAAATTGATAATAATGGCGTCAATAAATTATCTAACAC
     AGTTGGAATAACTTTACGTAACCATTTTTCAATCGTAGCTAAAATATATGT
     TGCTACGAGCATAGGTAATACTTGTCCCTGATAACCTACTTCATTAATATG
     TAGCCCAAAGACATCCCAGTGTGGAATAGCTTTTCCTTCTTCTAAAGCTTT
     TGGAAAATCATATGCACTCATCAATCCAGGATGAACAAGTATCATACCTAG
 45  AGCTGCACCTAAATAAGGATTACCACCAAATCGCTTAGCTGCACTAAAACC
     AATAAGTATTGGTAATAATGTAAAAGGTGCATTTGCAAATATATTAATCAT
     ATCAGCCAAGCCAGAAAATTGACTATGTACATCTATGATAGATTTACCATC
     ATAAAACAAATCTTTAGCTGTGAAAATATTATTTAATCCCATTAACAAACC
     ACCAGCTACAATAGCCGGAATAATAGGAACAAAGATATCAGATAACATTTT
 50  CACAAATTTTTGGAATGGATTCATATGCTTAGATGATTTATCTTTAACTTC
     TGAAGT
```

     Sequence 3465
 55 step.1003e10.cons.ok
     CTGATGCAAAGCTCATTGAACACTCACCCTTCCTAACCCTTTTTTCTATTC
     TTATTTATAGTATCGCCAATGAGGTGTAAAGTTATTTTCTAATTCTAAACA
     CACACTTAAGCTATCATATCATAACTACGAATTCGCTTATAATTATTTCTT
     TTTATCACTAGGAGTAAAGCGAATTGTACTTGTTAATTCTAAGGCAAGTTC

```
      ATAAATCATTTTTGATAATACTTTTGTGTTATGTCTGACATAGTGTTCATT
      AGATATTTCAACTAAATTAGATGCCGTTAAAACTCTAATTCCACTATCTTT
      TAATTGTTCTTTATGTACTGCTACTGGTTTCGAATTCTTTTCTTCATATCG
      TTGTAAAACATCTTTACTATAGGATTCTGAGCTACATATGACAAAATCAAT
    5 AAATGGTTCACCAACTTGTCGAGTAAGTGCATCAATATGCTCTTTGACATC
      ATAATTATCAGTCTCGCCTGGTTGTGTCATAACATTAGATACATAAAGTTT
      TGGAGCAGATGTACGTAATAATGCTTCTGAAATACCTTTGACACATAAATT
      TGATATAACACTTGTATATAATGAACCTGGTCCTAAGACAATTAAATCTGC
      TTGTTCTAAAGCTTCTATCGCTTCATTCATTGGTTCAACATCACTTGGTTC
   10 TAAAAACACACGATCTATTTTTTTATGTGTTTTAGGTATATTAGTTTCTCC
      ATGTACAATTTCACCGTCTTCCATCACCGCGTTGAGTTGCACACTTGCGTT
      TGTTGAAGGGATGACTTGGCCTTTAATATTTAAAACTTTGCTTAACTCTTT
      AATAGCGTGTCCAAAATCATTAGTAATGTTAGTCATTCCAGCAATAACTAA
      ATTACCTAATGAATGCCCATCTACTTGATTTTCACCAAAACGGTACTGGAA
   15 CAATTGAGTTAATATCGATTCTGAGTCACTTAAAGCAGCAATGACATTACG
      AATATCACCAGGCGCTGGAATATCCATGACATCTCTAATTTTCCCCGTGCT
      CCCACCATTGTCCGCTACAGTAACAATGGCAGTAATGTCTATTGGAAATTC
      TCTAAGGCCTCTAGCAAGGACAGAAAGTCCAGTGCCACCACCAATAAGTAC
      TACATTCATTTGTTTCATTATCTCTCTCCACTTTCAATATGCGCATCTCTA
   20 TGATGCACATAAACATTATATTCAAAAATCTCGTTAAGATATTCAGCTAAA
      CGTTTAGCTAATGCGACTGATCGATGTTGTCCACCCGTACAACCTATAGCA
      ATAACCAATTGCGATTTACCTTCTTTTTTTGTAGCCAGGAATCATAAATTTT
      AATAAATCTGTTAATTTATCAAAAAATATTTGGGTTTCTTTCCACTTCATA
      ACGTAATTGTACACTGGCTCATCTAAACCAGTAAATGGACGCAATTCCTCT
   25 ACATAGTAGGGATTAGGTAGAAATCTGACATCAAAAACTAAATCAGCATCC
      ATTTGTATACCATGCTTGAAACCGAAACTTGTCACGTTGATTGTAAATGTT
      TCAAAGTTTTCATCTAAATAAAACTTTGAAATGCGTTGCTTTAATTCTTTA
      GGTTTTAATTTTGTTGTATCAATCACGTAATTAGCGATACTTCGGATTTCT
      GATAGATGTTCACGTTCCTCATTTATTGCATCTATTAATGATCTTTGTCCT
   30 TGTTCATTTAGTGGGTGCGCTCTTCTTGATTCTTTATAACGTGAAATAATT
      TTTTCAGTTTTAGCTTCTAAAAACATAACATCTAAAATCACGTCATTACGA
      CTTTTAATAATATCAATTTCTTTAACTAGAGATTTAAATAATTCCTTACCT
      CTTAAATCTATTGCAATTGCTACTTTTTGCAATGAAGGTCAATGTGTATTA
      CTGGTATGCAATTTGGCAATGAGGTAGCTATGCATGTGAGCAGAGGTGCAG
   35 TCTTCGGTATGATAGGAGTTTTATGTAGTATACTAGCAACCTGGGGATTAC
      TACAAGCTACACATATGTGGCTATTGAGTATTATTGGTGGTTTTGTAGCTT
      GGTTTGTAAGTGCATTGATTATTTTTGAAATTGTGGAATTCATAGCACATA
      AAAGAAGGGATAAACATGGCTGGAAGACCAAAAGATCCAACAATCAATAAA
      AAAATTTATACTGAGATTCAACGTTTATTGGAAACGACTCATTTTAGAGAT
   40 ATTACTATAGATCAAATTTCTGAGAATACTGGTATTTCTAAAGCAACTATT
      TATAGACGTTGGAAGGATAAATCTTCAATTATCATGTCCGCGTTTATTGAA
      CAATCTCAATATATTGCGATTCATAATCAAGATAATTTATATGATGATTTA
      TTCCAGTTTTTAGTAAAAATAAAAGATATCTATAAAACAAAACTAGGTAGT
      GCTGTGATTGAAATATTAATTAGTCATCAACAAATGGAAGCTAGAGAAACT
   45 TTTATGACTAATTACTTTAATCATAATCGCAAAGTTTTAAAAGAGATTGTT
      CGTAAGCACATACAAGAGGAAGAACAAGATTTGTTTATTGATTTAATCTTC
      TCACCCATCTATTTTAATATATTAATTAAACCTGAAACTCTGGATGAAAAT
      TACATTAAAAAGATGTTGAATCAAGTTTTACGTATATATCATTGATATGAA
      AAACGCTCCTTTCGAGGAGCGTTTTTTGTTTTTCATGTAACAGGTTAGCCT
   50 AAGGTCAGTCTATTAAAACCTTTAACCATACTATGTCAAAGCATTTTGACA
      ATGATAAAGGTGACATCGCAAAAACTATTTGATAGCTTTTATAGAGAAAAC
      CTGATTTGATTTAATAATAATATTAAATGATTTCAAATCATAATTAAGGAG
      TTTAACGTAATGAATGTTAGCAATCAAATCAAAAAATTTAGAGAAAGAGAT
      GGGTATTCACAAGAATTTTTGGCTGAGAAAATGTTTGTTTCAAGACAAACC
   55 ATTTCAAATTGGGAAAATGATAAAAGTTATCCGGATATTCATAATTTACTT
      ATTATGTGTCAACTCTTCAAGGTATCTCTAGACGAACTAGTGGAGGATGAT
      TTGAAAAACGTGCAAATCAAAAATATTAAAAAAGAATTAGACTTTTGGACA
      TGGATGATGATAATTCCAATGGTATTAGGAAGTGTTCTCATAGGTCCTATG
      ATTCATTACTTTCATTGGTTAGGAATAGGAATAACATATTTAGTATTTTTA
```

ATAGGTATCATTGCTTCTACAAAATTAGAAAAAATTAAAATCAAAAATGAT
ATTGAAACATACGACAGGATATCTTTAGAGCACAGTGGCGATGATATC

5   Sequence 3466
step.1003e11.cons.ok
TATAAATTTAAATGTACATTAGCAAATGCATCGTTATCAAGTCCGTCTACA
GGGTCACCTACTACGACAAGTCCTGTTGCGGTATTATCTGATATTAAATCT
GATAAAGTAAAATTTGGAAACCAATTTTTATATCTTGCATCTGGAATTTCA
10   ATGCCTGGCGCGATACGGGTATGATATCTAATTGTTTCTAAATCAGCATCA
TAAGGTAAGTCTTCCTGCACTATAAAGATAATTTCTGGTTCTAGTAATGGT
GAAAATAATTCTGAAAGAGAGACTGAGGCACCATCATTAACAATTTGGTTA
GATAAGAGTGTTCCATATGCAGGTTCGTTAGTGTTAGCAATTGCTTGCGTT
GCCTTGCTAGTCATACTAACTTTATACCCTGTAACAGTCGAACGGTCTTTG
15   AAAGTTAATTGGTCAATTAGTTCTTCTTGTACATGATATGCTTCTTCTTCG
TTTAATTGATAGTTCTCACTAATGAAATCGATGGGTTTTTTATATCTATAA
GCTTTAAATAAAACTTTAGCAACTTCTTTATTTGTTAGTGTCATATTGTAC
CCCTCCCCTAAATTATCTGAAAATTTATACTATTATGATAACTTTTAAATG
TGAAGTGTGCAAATTGATTTAAATATACTTCATTTAGGGTATAAAATAATT
20   AATAATGTAAATTTTGATAGGAGTGGGTATAGTGTCAGATTATGTTTATGA
GTTGATGAAACAACATCATTCAGTTAGAAAATTTAAGAATCAACCACTTGG
TTCTGAAACGGTAGAAAAATTAGTAGAGGCGGGACAGAGTGCTTCTACATC
CAGTTATCTTCAAACTTATTCTATTATTGGTGTTGAAGATCCAAGCATTAA
AGCGCGTTTAAAGGAAGTGTCAGGTCAGCCTTATGTTTTAGATAATGGTTA
25   TTTATTTGTATTTGTTTTAGATTATTATCGTCATCATTTAGTAGATGAAGT
TGCGGCGTCAAATATGGAGACATCATATGGTTCTGCAGAAGGACTATTAGT
AGGTACAATAGATGTTGCATTAGTTGCGCAAAACATGGCAGTTGCTGCCGA
AGATATGGGGTATGGAATTGTTTATTTAGGGTCATTGCGTAATGATGTTGC
GCGAGTGCGTGAAATTTTAAATTTACCTGATTATACGTTTCCGTTATTTGG
30   TATGGCAGTAGGTGAACCTTCTGATGAAGAAAATGGGTCACCTAAACCGCG
CTTGCCATTTAAACATATTTTTTCATAAAGACCAGTATGATGCGAATCAGCA
TCAACAACGTAAAGAATTGGAAGCATACGACCAAGTAGTGAGTGAATATTA
TAAAGAACGTACTCACGGTGTGCGTACAGAAAATTGGTCACAACAAATAGA
AACATTTCTAGGACGTAAAACACGTTTAGATATGTTAGATGAATTGAAAAA
35   AGCAGGATTTATTCAAAGATAAAAAGAACCTGAGACAGAAATACATGTCTC
AGGCTAGGGTGGGGCGATATTTTCAACACGAATGTTGCCCCGCTCTTTTTA
TAATTGTATGTCGTAATGATCAATGACCACTTTTCTTACATTTAATTTGCA
GATAAATCGCCGTAATCATTTGAGTCAAATGTTTTTTTGTCCCAATGATTC
GTTAAGCGTGCGGTACCTGTTCCAGCTAGCATTGAGTCATTAACGTTAAGT
40   GCTGTACGACCCATATCGATGAGAGGTTCGATAGATATCAGTACCCCTGCG
AGAGCAACTGGTAGATTAAGTGTAGATAATACGAGTATTGATGCGAATGTT
GCCCCGCCACCTACACCTGCAACGCCAAATGAACTTATAATAACAACAGCA
ATAAGTGTAACAACAAATTGGAAGTCAATTTCTACATTTGCTACTGGTGCC
ACCATAACTGCTAGCATAGCAGGATAGATTCCTGCACAGCCATTTTGCCCT
45   ATGGATAAACCAAAAGTTGCAGAGAAGTTTGCAATTCCCTCAGGTACACCT
AAACGTTTTGTTTGCGTTTGAACATTTAACGGTAATGCACCTGCACTTGAA
CGTGAAGTAAATGCAAAGATTAGTACTTCTATTGTCTTTTTCACGTATTTA
ACGGGATTGATACCTAAGACACTCAGTATAATTAAATGGATAATATACATT
GTGATTAGAGCTGCGTATGAAGCAATTAAGAATTTACCTAACGTCCAAATT
50   GCAGAAAAATCACTTGTCGCAAGAGTAGAAGCCATAATAGCTAAAATGCCA
TAAGGCGTTAATCGTAAAACAAAAGTTACGATAGCCATAACGATAGAATAG
ATTGCTTCTATACCACGTTTAAGTAAGCTTCCATGTTCCGGCTGTTTTCTT
GCAACTCTAAGATAAGCAAAGCCCACAAACGTTGCAAAAATAACAACTGCA
ATTGTCGAAGTTGTACGTTGTCCTGTGAAATCTAAAAATGGATTGCTTGGG
55   AATACTTCGAGAATTTGTTGTGGTAAAGTGTTTGCAGTTAAATCTTTGGCT
TGTTTTGAAATTTCTGTACCACGTGAATGTTCTGCACTACCTAAATCAATA
GACGATGCATCTAAACCAAAGATCAAAGCGTAAAAAATTCCAACGATAGCT
GCAATGGCTACAGTACCAATTAAAAAACATAAAAATATAAGAACCGATCTTA
GCGAATTTTTCACCAATTTGTATTTTGCTAAAAGCGGCAACAATTGAAATG

AATATTAGTGGCATGACAATCATTTGTAATAGTGCAACATAACCATCTCCA
ACAATACTAAACCAGTCTGTTGATTGTTCGAGAGTTTTAGACTCTGCACCA
TATATGAGGTGTAATACAATGCCAAATACGATACCTATTGCTAAAGCGGTA
AAGACACGTTTAGGGAATGAGACGTGTTTTCTTGCCATAATATTTAACATT
5   ATTAGGAAAACAAGTAACACGATGACATTAATTACTGTAAAAAGTATTTCC
ATATCGATA

Sequence 3467
10   step.1003f01.cons.ok
TATAAGATAATATAAACTAAGCCTGAGACACCGATGAGTAATCTCAGGCTT
TTTTACGTATTGAAAGTGATTGAGTAATTAAATAGGGTGGATATCAAAAAT
GATGTTGGTCGATATGTTAGGTTCTTTAAGAGAATGCAATGCCAGGCTCAA
TATCGACGTAGTTATGAAGCGGTGTGTATTTTAAATGGCCGTTAGTATCAT
15   GTATAAACGGACTATTGGAACTCAACTTTTAAAATTGATATAGTATAATAT
AAAGGAAATTTAGAGGACGTGACGTTTCGATATTTGAATAACTGTATTAAG
GAATATCGAAACTTTAAAATATCTACGCAATTTAGGATAAAAAATCAACAT
TGAGTGAAAGCCAACTTAGTCGAAGCATCTCAATAAAGCGTAAGACACTCG
AGTTCCCCCACACCCCTATAAACAAAATAACTATGGTACATATGAGAAAGG
20   ATGAGTCGCTTGATACGTATGAGTGTATTAGCAAGTGGTAGTACAGGTAAC
GCCACTTACGTAGAAAGTGAGAAAGGTAGCCTACTCGTCGATGTAGGTTTA
ACTGGCAAGAAGATGGAAGAACTTTTCAGCCAAATCGACAGAAACATTAAG
GATTTAAACGGAATTTTAGTGACACATGAACACATCGACCATATTAAAGGT
CTTGGTGTTTTAGCACGTAAATATAAACTTCCGATTTACGCGAATGAGAAT
25   ACATGGAAAGCGATAGAGAAGAAAGATAGCCGCATTCCAATGGATCAGAAA
TTTATCTTTAATCCATATGAAACGAAATCTCTTGCAGGATTTGATATAGAA
TCATTTAACGTGTCACATGACGCGATTGATCCACAATTCTACATCTTCCAC
AATAACTATAAGAAATTTACGATGATAACTGACACTGGTTACGTTTCAGAT
CGTATGAAAGGTATGATTCAAGGTAGTGATGTCTTTATGTTTGAAAGTAAT
30   CACGATGTCGATATGTTACGCATGTGTCGCTATCCATGGAAGACGAAACAA
CGTATTTTAAGTGATATGGGTCACGTATCCAATGAAGACGCGGGTCTTGCG
ATGAGTGATGTCATTACAGGTAATACGAAACGTATATACCTCTCTCATTTG
TCACAAGACAATAATATGAAAGACCTCGCACGCATGAGTGTTGGACAAGTG
CTCAACGAACACGATATCGATACAGAGAAAGAAGTATTGCTTTGCGATACC
35   GATAAAGCACAAGCCACACCGATTTATACACTATAATTGAATAGAGGTAAG
ACAAACGACCTACTTAAGTATTACGCTTAAGTAGGTTATTTTTTTGGAAAA
GTGATCAGGAGAATTGCAGGGAATGATTATGTGTTAGTTTAAACAGGCTTG
CGCATGAACGATTGATTAAGGAATTTTTTTGAAATGAAAGCGCATCATGGA
AGGGGGAGAACGATGTCTTACACGCTTTGTGATTTATATATAATAACATTC
40   TTTTACTTTTAAACTGCTTAAATACCGTTCCTTCTGCGACATTTGCACGCT
GCGCAATTTCTTTAGTACTCGTTTTATCAAATCCTTGTTCACTGAATAATG
CAATAGCACTTTCTATGACACGTCTTTGTTTATCGCTTAAATATTCAAGTT
GAGGCACAATGGTTTCAACTAATGACTTAATATCTTGGTTCATATTGACCT
CCTAAACTTTTCTATATCTTTTCATGCCTATAATATTTAAAATCAATAAGA
45   AAATGAAAAATGCGAATAACACAATTAAATAAATGTAAATATCGTTGAATC
CATAACCCTTGATCATAATATTTTGCATCGTTTGGCCGGTATAGAATAACG
GCATGATATGTGAAAAGTATTGTAATCCTTTATTCATTGATTCAATTGGTA
TAATGCCTGCAAATAGTACTTGTGGCACTATGACTAATGGTATAAATTGAA
TCATTTGGAATTCTGAGGAAGCAAAGGTAGATAATAATATACCGAATGTCA
50   CAGCGACAAGCGCTGTTAATATTGCCGTTAATAGTACGAACCATATCGAAC
CTACTAAGTCTATATGCAGAATATAAATTGCATATAATACGACAACTATTG
TTTGGATAACGCTAAAACTACCATAACCGAAAACATAACCAAAAATAATTT
CACTTCTTTTTATTGGAGAGGCAAGTAAACGTTCTAATGTGCCAGAAGTAC
GCTCTTTTAATAAGCCAATGCCAGAAATTAAAAACGTAAAGAAAAAGACAA
55   AAAATCCAATTAAAATAGGATTTATCATATCAAAATACGTAGAATCTGATG
AACCATATAAATAGTGCGTTGTTAGTTTATATGGTTTAGCCATATCTTGAT
GAGGCGTGTCTCCCCCATCCCCGGGCATTTTTTGTTGTATTTTAGTTAACG
CTTGATGCAATTTATTAGTATTATCTTTCATGGCATTCATGTTATGACTCA
TTAACCATTTTTGATTTGCACCTGTTAGTTCTCCTGCTTGTGTAGGATTAT

CGTTAGCATAAGTCACTGATACTTTTTGACCATCACTGTGCAAAAATCCTG
TTAATTTGTCGTCTTTAATTTTATCACTTATATCATTGTCATTTTTATAAT
GTTTAACGTGAATATCTTTATCATGTAATTCAGTCATTAATGAATCTGGTA
CATCGTGAACCCCAACTGTTACGCCATTCGTATTATCGGCAACTGTAAAAA
5 TGTAATATAGTAGCGACAATAGTAATATAGGTGCAATAAGTAACAGTGCCA
ATGTTCTTACATCTCGAATCGTCTGTTTGAAAATACGTCCTGCTATGTGCA
ACGAATTCATTTTATCCCTCCATATTCAGAAATACGCCTTCAATCGTATCT
GTGTGATATTGTTTTTTCACTTCATCTGGTGTACCCGTTGCCAATATCTTT
CCTTGATTCATTAATACGAGCTTATCACAGCGTGTAGCTTCATCTAATACA
10 TGCGTTGTCACTATAATGCATTTATCTTCAGCCTTTGCTTTAGTTAAATCC
TTCCAAATCTTTAGAGCACAGTGGCGATGATAT

Sequence 3468
15 step.1003f02.cons.ok
TACAGAAAATAAGACCTCTAAAATGCTAACTAGTTAGTATTTTAGAAGGCT
TTTTTATAAAAAAGAGAAACATCGAAGGGAAGATTCATTCGACATTTCTCC
ATAATGATATTAATTAAGCTTTTATGCTTAGTTGTATTGTTAATTATTTAC
CTTCTCCGTAGATGTCATTCCATTGCTCTCTTTTATCTAAAAATGCTTGAG
20 CAATTGTTTTAGCACCTTCTAAAGAATGACTTGCAGCCCAACCACATTGAA
CCTCATTACAAGCTGGGACTTCGCTAGCATTTAACACATCATGCAATGTTT
GATCGATAATATTTAATACGTCATCGTAGTCGTCATGATTAATAAATGAAA
CATAGAATCCAGTTTGACAACCCATAGGACTTAAATCTACTACTTTATCAG
TATGATTTCTAATGTTTTCTGCCATTAAATGCTCTAAGGAATGTAGACCAG
25 GCATATCCATATGTTCCTTATTGGGCTGTTTGAAACGAATGTCATATTTGT
GTATGACATCACCATTAAGACCTTCCATAGTCCCGGCTAGACGAATAAAAG
GTGCAACAACCTTAGTATGGTCTAAATTAAAGCTTTCTACATTCATTTTAG
TCATGCATTTATCCTCCTTATTGCTCTACAGCAATTTTTTGTTTAATTCAA
TAAATAGATTGTAACAATACTCTAATACATTTACAATTTTTCTAAACTAAT
30 AGACTGTTGTCTTTTAAATTGTAAGAAAATTGTGTAAAAATTAATGACAAC
CAATATTCAATCCGATAAAATAAGTAATAATTAAAAAAGAATGATTGATAC
GCAAAGTTTCAAAGGGAGTTAGATACAATGCATAATAAACAAAAGATATTA
GATTTTATAGAAAATAATAAATATGATTATGTTGAAATAAGTCATCGTATT
CATGAACGCCCTGAATTAGGCAATGAAGAAATTTTTGCATCGAGAACATTA
35 ATTGACCAATTAAGAGCAAATCGATTCGAAATCGAAACGGATATTGCAGGA
CATGCAACAGGATTTATAGCAACGTATGATTCTGATATGACTGGACCGGTT
ATAGGATTTCTAGCTGAATATGATGCTTTACCTGGTCTTGGTCACGCATGC
GGGCATAATATTATTGGTACTGCTAGCGTACTTGCTGCAGTAGCACTAAAA
GAAGTCGTCGATGAAATTGGTGGTAAAGTAGTCGTTTTGGGATGTCCTGCT
40 GAAGAAGGTGGGGAAAATGGCTCCGCAAAAGCTTCTTATGTTAAAGCAGGT
GTCATTGATGAAATTGATGTAGCATTGATGATTCATCCTGGAAATGAAACT
TATCGTACAATTAATACTTTAGCTGTGGATGTTCTAGATATTAAATTCTAT
GGACGTAGTGCGCATGCATCTGAAAATGCAGATGAAGCATTAAACGCTTTA
GATGCAATGATTTCATATATTAATGGTATAGCACAGTTAAGGCAACACATT
45 AAAAAAGGACAACGAGTTCACGGGGGTTATTTTAGACGGTGGTAAAGCGGCT
AATATTATACCTGATTTTACACATGCGAGATTTTACACTCGAGCTACTTCG
CGGAGAGAACTTGATGTTTTAACTGAAAAAGTAAACCAAATTGCAAGAGGT
GCTGCTATTCAAACTGGGTGTGATTTTGAATTTGGTCCTATCCAGAATGGT
GTAAACGAATTTATCAAAGCACCTAAACTTGATGATTTATTTGAAAAATAT
50 GCAACTGAATTAGGAGAAGAAGTGATTGATGATGATTTTGGCTATGGATCT
ACAGATACAGGTAATGTAAGTCATGTTGTACCAACAATACATCCACATATT
AAAATTGGTTCTCGAAATCTTGTTAGGACATACCCACCGCTTTAGAGAAGC
GGCTGCAAGTTTACAAGGTGATCAGGCACTCATTCGAGGTGCAAAAATTTT
AGCATTAATGGGACTAGAATTAATCGAAAATAAACCGTTGTTAGACGAAAT
55 AATTGAAGAGCATACGCATATAAAGGGGCATGTTAAGTAATGAATAATAAA
GTAAATGGTCCGTTTCTCACACTTAGTGATTTGTATAATGATGACATTGTT
TACACATCTCGACCTTCATATGTGTCGAACCCATGGTTGAAACCAGATGAA
CACCAATCAAATTTCTTAACTGGAAGAGAATTACTTATAGCTAATCGTTTT
CCTGTCATTGTTCATGAGGCAAGTGTTACAAACAAATTAGAACAACTTTTT

AATATAGTGGGTAGGGAAATTCCACCACATACTTATAAATTTAAAGATCAA
GAAACATATGAAAGCTTGATAAGGAATTTAGCTCTTCATCAAGGTAAAAAA
ATCTACTTTCAATATATTCATGATGAAGATATTTTACCTAAAGAATATTAT
GCACTTGATAAAGATGTTTTTGTTGCTCTTAATAATAAAGCACGAATTCCA
GAATGGACTAATAACAAATATCTACCACAAAGAGAAATTGTCTCAATTAAA
GATTTTGAATCTCATATTCAAGCATGGTCGTATCCATTTGTCATAAAACCA
GGCGATGATTTACCTACAGCAGGAGGATATGGTGTTATGATTTGTTATAAT
GATACAGATTTAGCTAAAGCTATCACACGCATCAACAATGCATCAGCAGAG
ACTGAAAATTTAATCATTGAACAAAAAATTAATGCAGTGAATAACTATTGT
GTACAATTTGCTTATTCAGATGATATTGGTATCAAATACTTAGGAACAGCG
CAACAGTTAACTAATGACTATGGATTTTACAACGGAAATGAAAATGTTAAT
GATGTGCCTCAGAATGTAATAGACGCTGGTAGAGAGATTATGGAAATAGGC
GTAAGCAAAGGTTTTTTTGGTGTAGCAGGTTTTGACTTACTAGTAGATGAT
AATAATGATGTTTATGCGATTGATTTAAACTTTAGGCAAAACGGATCAACG
AGTATGCTACTTTTAGCAAAAGATTTAACTCATGGATATCATAAATTTTAC
AGTTACTTTTCTAATGGAGATAATACAAAATTCTATAATGCTATTTTAAAA
TACGTAGAATTAGGTGTACTTTATCCACTTTCCTATTACGATGGAGATTGG
TATGGAAAGAATCAAGTTAATTCTAGATTTGGCTGCATTTGGCATGGGGAA
AATAAAGAATTAATTAATCGATATGAACAACAATTTATATTGGAAGCTGGA
TTATA

Sequence 3469
step.1003f03.cons.ok
ATGAAGCAGCCGCTGGATCAAGATTTTCGGGAACTTTAACTGCATAATTCG
CATCAACAATGACTTCTTCAGCCATTGCCCCATCAACAGTATATCCGGCAT
TAAGGACATCCCTGCATAATGTTTCTCTACCAGTTGTGCAATACTCACAAT
TACCACATGCTGCATACATCCATGCTATAGATACACGATCCCCTACCTTTA
ATGAATCTACGTTATCAGCAATTTTAATGACACGTCCAATTCCTTCATGTC
CTAATGTTACACCTGTTACGTCACCGAAATCTGCATTTTTAACATGTAAAT
CAGTATGACATACACCACAATATTCAGTCTTTACTAATGCTTCTCCAGGTC
TTAATTCTCTTAAACGTTTTTCTTCAATACTTACTCTATGATCTTTGGTTA
CTACAGCTGCTTTCATAAAAAGCCCTCCCTAGTATTAAAGTGTAATTCTAT
TACTTCCACCCTTATTTTACACTCTTTTGAAATCGTTTTCCAAAAGTCCTT
TCATTTATACTGAATTTTTTTGTAAATTGTATATAGAGCGACATAATATCGT
CACTTTTATAGCATATTTTATGATTTATTCATATATTTTATTTTTTAACTA
TATATATATATCATGTCAGTGAAAAGTCATCGCTCTTAATCACACCTTATA
ATTTTGGGGATGAAGGTGTTTATTTTAAAAATCTTAATTGGAGGGGATAAA
AGTGAAGACGCAAGTTTTGTACGGAAAATCTGAAATTGAAATCAATGTGCC
TGATGATAGCACAATCATAGAGCCTCAAAACATCGACGCTATTCAAGATTA
TGAATCAACAATTAAAAATGTATTAAGAAACCCCACAAATTCTAAACCTTT
AAAAGAAATGGTAAACAGTAATGATATTGTTTCTATTGTTATTAGTGATAT
TACACGTCCAACGCCCAACCATATTCTTGTACCTTTACTAATTGAGGAATT
AAATCATGTTCCTCGTGAGAATTTCGTAATTATTAATGGTACAGGGACTCA
TCGAGATCAAACGCGAGATGAATTGATTCAAATGTTAGGTGAAGATATTGT
AAATTCAGTAAAAATCGTTCACAATCATTGCTCAGAAAAAGAAAGTCTAGC
TAAAGTGGGACACAGTCAATATGGATGTGATGTTTATTTAAACAAAGCATA
TGTAGAATCCGATTTTAAAATTGTAACAGGTTTTATTGAACCACACTTTTT
CGCCGGATTTTCAGGTGGACCTAAAGGGATAATGCCTGGAATTGCAGGTTT
AGAAACAATTCAAACATTTCATAATGCAAAAATGATTGGCGATCCGAGATC
AACGTGGGGAAATTTAGAAGACAATCCAGTTCAAGATATGGCACGGGAAGT
TAACCGTATGTGTAAACCTGACTTTTTTACTTAATGTTGCATTGAATAAAAG
TAAAGAAATTACTGCAGCATTTGCTGGTGAAATCTTAGATACACACAAAGA
AGGATGCGCATATGTAAAAGATCATGCAATGTTTAAATGTGAGCAACGCTT
TGATATTGTTATCGCATCAAATTCTGGCTATCCTTTAGATCAAAATTTATA
TCAAACAGTTAAAGGGATGAGTGCAGCGAGTAAAGTTGTTAAAAAAGACGG
TCATATTATTATGGTATCTGAGTGTGCAGATGGCTTTCCTGATCATGGTAA
GTTTGCCGAAATTTTCAAAATGGCAGACACACCTCAAGGTATTTTAGAACT
TATTCACAATCCAAACTTTAAGGAAGTTGACCAATGGCAAGTACAAAAACA

992

```
AGCAAGTATTCAAACTTTTGCCAATGTGCATGTTTATTCAGAACTTACTGA
CCAACAACTTAAAGACTCGATGTTAATCCCAACCTCTAACATTGAACATAC
AATACAAGAATTAGAACATCGATATGGCCGTAAATTAACCATTGGTGTTAT
GCCACAAGGTCCTTTAACAATACCGTACGTAGAAGATAAAGAATAATAAAT
GGAGGATGGAAACTATGACTGAATTAAATACTAAAGAAATGAAATTAGACG
CACTATTGAAAGACATGCAGAGTGTAGTAATTGCCTTCTCAGGTGGAGTAG
ATAGTAGCTTGTTACTGAAAAAAGCGATTGATATTTTAGGTGTTAACTATG
TTAAACCTGTTGTAGTAAAATCAGAATTATTTAGAAATGAAGAGTTTGAAC
TAGCGCTTAAACTTGGACAAAGTCTAGGTGTTGAAGTATTAGAAACTGAAA
TGTCTGAACTTCAAGATGCGAATATCGTTAAAAATACGCCTGAAAGTTGGT
ACTATAGCAAGCGCTTGATGTATAGTCAACTTGAGAATATTAAGAATAAAC
TAGGATTTAATTATGTGCTAGATGGTATGATTATGGATGACTTAGATGATT
TTCGTCCCGGATTAAAAGCAAGAGACGACTTTGGTGTTCGTAGCGTTTTAC
AAGAAGCAAAACTATCTTTAGAGCACAGTGGCGATGATATC
```

Sequence 3470
step.1003f04.cons.ok

```
AGGTGAAGATGTATTCTTAGGATTTGACTGTGCTTCTTCTGAATTCTATGA
AAATGGTGTTTATGATTACACTAAATTCGAAGGTGAACACGGTGCTAAACG
TAGTGCAGCAGAGCAAGTTGACTACTTAGAAGAATTAATTGGTAAATATCC
AATCATCACTATTGAAGATGGTATGGATGAAAACGATTGGGAAGGTTGGAA
ACAATTAACTGATCGTATCGGTGATAAAGTTCAATTAGTTGGTGATGATTT
ATTCGTAACTAACACTGAAATTTTATCTAAAGGTATCGAACAAGGTATTGG
TAACTCAATCTTAATCAAAGTAAACCAAATCGGTACATTAACTGAAACATT
CGATGCTATTGAAATGGCTCAAAAAGCTGGATATACTGCGGTTGTATCTCA
CCGTTCTGGTGAAÀCTGAAGATACTACAATTGCTGATATCGCAGTTGCTAC
AAATGCAGGCCAAATTAAAACAGGTTCATTATCTAGAACTGACCGTATTGC
TAAATACAATCAATTATTACGTATTGAAGATGAATTATACGAAACAGCTAA
ATTTGAAGGAATTAAATCTTTCTACAATTTAGATAAATAATATCTCTTTAA
TTTTTATAAGAGTTCACGCTTATTATTTAATATCTTAATTGAATTGTCCGT
TGCAGAGTTGAAGCTGCAACGGGCTTTTTTAGTTAAGTACATTTTAAATTA
TTAATGTAAACTGATATAATGATGTGAGTAATTTTGTAAAAGGAGTACAAT
TATGACAGACTCAAATGCTAAAGAAATAAGAACTGGACGTTTAATTGCGAT
AAGTTCATTAGTGTTTTGTATTTTACTTATCATACACCACTTTATTGTATT
AGATGAATCAACAGCTAAATCAATTTTATCTTTAGCTGGTCAAAAAACATC
AGATACAGCAGTGAAAAACATTTTAAATAGTGACCGATACACTGGAATTAT
GTATATTTTAGCTTACTTAGCAGGTACTGTTGCTTTCTGGAATCGCCATCC
ATATTTATGGTGGTTTATGTTTGCCGTATATATTTCTAATGCACTATTTAC
ACTCGTAAATCTTTACTTATTTATTCAAGGTATTTTAGATGTAAAAAATGT
ACTTGCAGTTTTACCAATTTTAATTGTAGTGATTGGATCTATAATTCTAGC
AATTTATATGCTAGTTGTTTCTATTACACGTAAAAGTACTTTCAATAGATA
GAAGTTTGATTACATCAATAAAGATGTGTTATAATTTCTTTCGTATATAAT
GAACGGAGGACAATTTTTATGCATACATTAATCATCGTTTTATTAATTATA
GATTGTATTGCATTAGTGACTGTTGTATTACTCCAAGAAGGTAAAAGTAAT
GGACTTTCAGGTGCTATTAGTGGTGGCGCTGAACAATTGTTTGGTAAACAA
AAACAACGTGGCGTCGATTTATTCTTGCATAGATTAACAATAATTTTAGCT
ATTTTATTCTTTGTACTTATGTTTTGTATTAGTTACTTAGGTATGTAGAGA
ATATTATATTTGTTTTAATGAAGTAAATGAATCTTAACCAAGACTGATAAA
TGATAGGGCAAATTAATCACATTTGCTATTAAGGTGATGCTGCCAACTTAC
TGATTTAGTGTATGATGGTGTTTTTGAGGTGCTCCAGTGGCTTCTGTTTCT
ATCAGCTGTCCCTCCTGTTCAGCTACTGACGGGGTGGTGCGTAACGGCAAA
AGCACCGCCGGACATCAGCGCTATCTCTGCTCTCACTGCCGTAAAACATGG
CAACTGCAGTTCACTTACACCGCTTCTCAACCCGGTACGCACCAGAAAATC
ATTGATATGGCCATGAATGGCGTTGGATGCCGGGCAACAGCCCGCATTATG
GGCGTTGGCCTCAACACGATTTTACGTCACTTAAAAAACTCAGGCCGCAGT
CGGTAACCTCGCGCATACAGCCGGGCAGTGACGTCATCGTCTGCGCGGAAA
TGGACGAACAGTGGGGCTATGTCGGGGCTAAATCGCGCCAGCGCTGGCTGT
TTTACGCGTATGACAGTCTCCGGAAGACGGTTGTTGCGCACGTATTCGGTG
```

AACGCACTATGGCGACGCTGGGGCGTCTTATGAGCCTGCTGTCACCCTTTG
ACGTGGTGATATGGATGACGGATGGCTGGCCGCTGTATGAATCCCGCCTGA
AGGGAAAGCTGCACGTAATCAGCAAGCGATATACGCAGCGAATTGAGCGGC
ATAACCTGAATCTGAGGCAGCACCTGGCACGGCTGGGACGGAAGTCGCTGT
5   CGTTCTCAAAATCGGTGGAGCTGCATGACAAAGTCATCGGGCATTATCTGA
ACATAAAACACTATCAATAAGTTGGAGTCATTACCTGCTATTAAGTATTAT
CAGCCTGAGTCCGACGATGTACTTGTCGGACTTTTTTATTTTATAATTGAG
ACAATCATGTTTTTAAATCATATGTTAAAGGGAATAATGATTTATCAAGTC
AATTTGATAATAATGAAGAGATTTTTAAATATATGAAATAAATTGAAATAC
10   TTAACTTATTTAGGATAAATTGAGTACTTACTATAAGGGAGAGGGATTTTA
ATCATGCAAATTAAACTACCAAAACCATTCTTTTTTGAAGAAGGGAAACGT
GCAGTGTTACTTCTTCACGGCTTTACAGGTAACTCTGCTGATGTAAGACAA
CTTGGGCGTTATCTTCAAAAAAAGGGCTATACATCTTATGCTCCACAATAT
GAAGGACATGCAGCGCCCCCAGAAGAAATATTAAAATCTAGCCCTTTTGTT
15   TGGTTTAAAGATGTTTTAGATGGTTATGATTATTTAGTAGATCAAGGTTAC
GAAGAAATAGCAGTAGCTGGTTTATCATTAGGTGGCGCCTTCGCATTAAAA
CTAAGTTTAAATCGTGATGTGAAGGGGATTATAACTATGTGTGCACCTATG
GAGAATAAAACAGAAGGTTCGATTTATGAAGGCTTTCTTGAATATGCACGT
AACTTTAAAAAATATGAAGGCAAAGATCAACAAACGATTGATCAAGAAATG
20   GAACAATTTCATCCAACTGAAACCCTGAGAGAACTGAGTGACACTCTAAAT
GGAGTTAAAGAACATGTCGATGAAGTAATTGATCCAATACTTGTCGTACAA
GCAGAACAAGATACAATGATTGATCCTCAATCAGCAAATTATATATATAAT
CATGTCGATTCTGATGAAAAAGAAATCAAATGGTATCAACATTCAGGTCAT
GTGATTACCATTGATAAAGAAAAAGAGAAAGTCTTTGAAGATGTATATCAA
25   TTTTTAGAATCATTGGAATGGACAGAGTAAAAAATAAAGATGTAGGAGAGA
AAGGAGGGGCATAATGAATTTAAAGCAATCCATCGAAGAAATGATAAAACA
ACCTGACTATGAACCCATGTCAGTATCTGACTTTCAAGATGCGTTAGGTTT
AAACAGTGCCGACTCATTTAGAGATTTAATTAAAAATACTCGTTGAATTAGA
ACAGTCTGGTTTAATTGAACGTACAAGAACAGACAGATATCAACGTAAACA
30   ATCCAACAAAACAAATTCAAAACTAATCAAAGGAACGTTAAGTCAAAATAA
AAAAGGCTTTGCTTTCTTAAGACCTGAAGATGACGAGATGGATGATATTTT
TATTCCACCAACTAAAATCAATAGAGCATTAGATGGAGATACTGTCATCGT
GGAAATTCAAAAATCTCGTGGAGAACATAAAGGTAAAATTGAAGGTGAAGT
AAAATCTATTGAAAAGCATTCAGTTACACAAGTTGTTGGAACGTATAGCGA
35   AGCAAAGCATTTTGGTTTCGTATTACCGGATGACAAACGTATTATGCAAGA
TATCTTTATACCTAAAGGACAAAATTTAGGTGCTGTAGATGGTCATAAAGT
ATTAGTACAAATTACGAAGTATGCCGATAGTACTGACAAT

40   Sequence 3471
     step.1003f07.cons.ok
     TAAGCAATAAAGTTCTTCAATTTTATGTAAAATATGGAGGAGAATTCCAAT
     TAAAACAAGGATTTAGTCCAGCGTTCACCGTAGAAAATAAAGATGCTATTG
     ATATAGGTTTTGAACAAACATTCTATGAAATAAATGTCGTTATTGCGGAAA
45   AAGATTTATGGTATTTTCAAGATGAAAAACTAACTGTAGATGCAATCGATC
     ATGAAGATGAGATCATCTACAAACGAAATTAATATAATTATACATATAAAA
     AAAGAAGCGATAAAAACTACTTCGCTTCTTTTTATAAATGATTAATGCCAA
     ACGTAACCTCAGCATCTTTACCTTCTTTATTTAACTGGTTAACTTTACTAT
     AAGTCTCATGCCACTCAGGAAAATATCTATCCAAGCGTATGGGTTTAAATG
50   TATCTGCTTTAATACAGATAAGTTCTGTTGAACCAGTAGTAGCTAACTCTC
     CATTTTCATTATAAACCTCATAACAATAAGTTGAACGTAATCTAGAATATT
     TTTCCACCCATGTTTTTACTGTCACCTTTTCAGGATAGTAAATTGATTTTT
     TATATTGCACTTTTAAATCTACAACTGGTGAAATAACACCTTGTTCTTCCA
     TAGAGGCATAACTGAAGCCAAGCTTTCTTATATAGTCTGTTCTCGCAACTT
55   CAAACCATGTTGCGTAGTTACCGTGATAGATAACCCCCATTTTATCGGTTT
     CTTGATATCTTGCTTCAATTTCAGTCAAACTGTATATCATTGGAATAATCC
     TTCTTTCAACTTTAACTTATTCTTTCAATTAAGGATAACATAAAGCAGTCA
     TGCCAACACAATAGTTAGCATGACTGCTTGTACTCAATAGTGAGTCAAAAT
     CATATTAAATTACTGCGCAAGTTTATTTCTTAGTACCATTTGTAAAATACC

```
     ACCATGACGATAATAATCTAATTCTACTAGTGAATCAAAACGTACAATAGC
     TTTGAAATTAATGATTTCACCATTTTCTTTTTTTGCAGTCACATTTACAAG
     ATCATGTGGCTGTACATCTTCATTAATATCTACAGATATTTCTTCTTTTCC
     ATCAAGACCCAGTGCTTCTGCAGATTCTCCTTGTTGGAATTGAAGCGGTAG
5    TACACCCATCATAACTAAGTTAGAGCGATGAATACGTTCATAGCTTTGTGC
     AATGACAGTTTTAACTCCTAATAAATTGGTACCTTTTGCAGCCCAGTCACG
     AGAAGATCCCATTCCATAGTCATTACCAGCTAAGACAACTAAGCCAGTTCC
     ATCTTCTTTATATTTCATTGCTGCATCATATATAGGCATTATTTCTCCGGT
     AGGCCAATATGTTGTAAATCCGCCTTCAGTACCTGGAGCAAGTTGGTTTTT
10   GATACGAATATTGGCAAATGTACCACGTACCATAACTTCGTGGTTACCGCG
     ACGGGAACCATAAGAGTTAAAGTTGCGAATTGCAACATCATGATCTAATAA
     GTATTTTCCTGCTGGTGTATCTTTACCGATAGCACCTGCTGGAGAAATATG
     GTCTGTTGTAACAGAATCACCAAATTTACCCATAACTCTCAAACTTTTAAG
     TGGTTCAATTTTACCCGGCTCTTTAGATAATCCTTGGAAAAATGTTGGATT
15   TTGAATATATGTTGAATTAGGATCGAAATCATATAATGGTTCATCGGTTAC
     ATCTATTTCATTCCACATTTCATTGTTATGATATACATTTTTTATATTCTTC
     TAAGAATAGTTCAGGTGTAACAACTTTATCTACAGTATCAGAAACTTCTTG
     TATTGAAGGCCAAATGTCTTGTAAATATACATCTTTACCATCTTTACCTTT
     ACCAATTGGTTCATTTTGTAAATCAATATCTACTGTACCAGCAAGCGCATA
20   AGCTACAACAAGTTGTGGTGAGGCTAAATAGTTTGCTTTCACTAATGGATG
     GATTCGCCCCTCAAAATTACGATTACCTGATAAAACTGAAGTTACTAATAA
     ATCTTCATCCGCAACTGCCTTTTCAATTTCAGGTAATAGTGGCCCTGAGTT
     ACCAATACATGTAGTACAACCATAACCAACAAGATTGAAACCTAAATCATC
     TAAATACTGTTGTAATCCAGAATCTCTTAAATATCCTGTAACAACTTTTGA
25   ACCTGGAGCAAGTGACGTCTTAACAAACTCTGGTACTTTCAATCCTTTTTC
     TACAGCTTTTTTTAGCAACTAAACCAGCACCTAACATAACATATGGATTAGA
     AGTGTTAGTACATGAGGTAATAGCAGCAATTGCTATATCTCCTGTTTTCAT
     TGTTGCTTTAGATCCATCATTAAAATTAATTTCTGCTTTTTTTATCAAATTC
     ACTTTGATCAAGTCCATGTCCTTGATTACCAGCAGGAGCAGTTACTGATTT
30   TTCAAATTCTTTTTTTCATATCACTTAAGAAAATTAAGTCTTGTGGACGCTT
     TGGACCAGAAAGTGATGCCTCTACTGTAGATAAGTCTAAATCAATAACATC
     TGTATATTCAGGATCCTCTTTTTCTACATCAAAGAACATATGGTTTTGTTG
     TAAATATTCTTTAACCAATTCAATATGTTCTTCGTCACGGCCTGTAAGTTT
     CATATATTTCAATGATTCTTCATCTACTGGGAAGAAACCACACGTTGCACC
35   ATATTCTGGAGCCATGTTAGCAATTGTAGCTCTGTCTGCTAATGGTAAATG
     TTGAACACCTGGACCGAAGAATTCAACAAATTTTCCAACTACACCTTTTTT
     ACGTAATTCTTCAGTCACACGTAAAGCTAAATCCGTAGCTGTTGAGCCTTG
     TGGTAAAGAGTGAGTTAAACGCACTCCGATAACTTCAGGAATTGGGAAATA
     TGATGGTTGTCCTAACATACCTGCTTCGGCTTCGATACCACCAACGCCCCA
40   ACCTAGAACACCAATACCATTAATCATTGTAGTATGTGAATCAGTACCTAC
     TAAAGTGTCAGGAAATGCTGTTTTTTCACCATCAACATCTCTTACATGTAC
     TACATTTGCTAAATACTCTAAGTTTACTTGATGGACAATACCTGTAGCAGG
     AGGTACTGCATTATAGTTATCAAAAGCTTTTGTTGCCCAGTTTAAAAATTG
     ATAACGTTCATAGTTACGTTCAAATTCTAATTTCATATTACGTTCTAATGC
45   TTCTGGATTAGCGTAACTATCAACTTGAACTGAATGGTCGATAACTAAATC
     CACAGGTACTTCTGGGTTGATTTTATTAATATCTCCACCAACATCATTCAT
     AGCTTTACGTAAAGAAGCCAAATCTACTACTGCTGGCACACCTGTAAAGTC
     TTGTAAAATAACTCTAGAAGGTTTGAATGGAACTTCACCTTCGTTACCTGC
     ATTTCCGAATTTACTTAATGCTTTGATATGATCATCTGTTATAACAAAATC
50   ATCCTCTTGTCGTAACACAGATTCTAACAATACGCGAATTGAGTATGGTAA
     TTTA
```

Sequence 3472
step.1003f08.cons.ok
AGCCACTGACCGTATTTACGGCCAGTGGCTTATTAAATATAATATAAAAAA
CCGAGCGTATAAACGCTCGGATATAATGATTTATTTTAATTGTAAATCAAC
TTCATTTATTTCTTTTAAAATTTCAGCTGTATCATAGCAATTTGTAGCGAT
ATTTGCATATTTCTCTGCTAAACGATAAGCATTAATGTATAAAATTAAACT

TTCTTTAGCGATATCTTCAGCGTCATCAGATTTGGCTGGATTTGAAGTTAC
AACTAATTCAGCAAATTTTTCAGGATCAATATTAATAGACATGAATGAATA
AACAACTCCTATATTTTTATTAGTTTCATATTAACACAATACTTAAATACA
AAAAAGTAAAGTCTTTTAATTTTCAGAATAATTAATTTGTTAATGAGACTT

5 TTATTATAGTACGGTATCCATTAAACTTAATTTATGAGAGGATGAACAGAG
AATGAGAATATTGTTTATAGGTGACATCGTTGGTAAAGTGGGCAGGAAAAT
GATTACTACTTATTTACCTAAAATTAAACAAACTTATCACCCAACAGTTTC
TATAGTAAACGCTGAAAATGCCGCACACGGTAAAGGATTAACAGAAAAAAT
TTACAAACAACTTTTGAGAGAAGGCGTGGATTTCATGACTATGGGTAATCA

10 TACATATGGTCAAAGAGAAATTTACGATTTTATTGATGATGCTCATCGAAT
GGTGAGACCTGCAAATTTTCCTGATGAAGCTCCAGGAACAGGTATGAGAAT
AATAAAAATTAACGATATTAAATTGGCTATTATTAATTTACAAGGCCGTTC
ATTTATGCAAGACATTGATGATCCATTTAAAAAGGCTGACCAGCTAATCGA
AGAAGCTCAAAAATCTACACCATATATATTTGTAGATTTTCATGCTGAAAC

15 TACATCTGAAAAAAATGCTATGGGTTGGTATTTAGATGGTAGAGTGAGCGC
TGTTGTTGGTACTCACACACATATTCAAACTTCTGATGATCGTATATTACC
TCATGGCACAGGATATATCACAGATGTCGGGATGACAGGTTATTACGATGG
TATTTTAGGTATCAATAGAGATGAAGTTATTCAACGTTTTATTACTAGTTT
GCCACAAAGGCATGTTGTTCCAGATGATGGGCGAGGCGTATTATCAGGAGT

20 TATCATAGATTTAGATAAAGAAGGTAAAACGACTCAAATAAAAAGACTGTT
AATAAATGAGGACCATCCTTTCCAAATTTAAAGGGATATAGTCCATTAGTA
TGAATTTAAACTTGTACCACTGTTGAAATAGTAAAACGGTGGTTTATTTTG
TTATCATAGAGTATGAATTAATACCACAGGAGGCATGTGATATGAAATCAC
AAATATCATGGAAAGTGGGCGGTCAGCAAGGCGAAGGTATTGAATCTACCG

25 GTGAAATCTTTGCTACTGCGATGAATAGAAAAGGTTATTTTTTGTATGGAT
ATAGACACTTTTCTAGTCGTATAAAAGGTGGCCATACTAATAATAAGATAA
GAGTTTCAAAATCGCCTGTGCATGCGATTAGTGATGATTTGGATATACTCA
TTGCTTTTGACCAGGAAACGATTGAATTAAATCATCATGAAATGAGAGAAG
ATAGTATTATAATTGCGGATGCTAAAGCAAAACCCCAAAAGCCAGAGAACT

30 GTGTGGCTCAATTAATTGAGTTACCATTCACTAGCACGGCAAAGGAACTTG
GAACAGCATTAATGAAGAATATGGTGGCAATTGGTGCGACATCTGCACTGA
TGGATTTAAATACATCAACTTTTGAAACTTTAATCGATAACATGTTTTCAA
AAAAAGGTAATAAAGTCGTTGATATGAATATACAAGCCCTTAATATGGGTT
ATGATTTAATGAAGCAACAAGTTACCAACGTTAATGGAGACTTTACATTAG

35 AGAATGGTAGCGGTCATCCTCATTTATATATGATAGGTAATGACGCAATCG
GATTAGGAGCAATAGCAGCTGGATCAAGATTTATGTCCGCTTATCCAATTA
CGCCAGCTTCTGAAATTATGGAATACATGATTGCCAATCTACCTAAAGTTG
ATGGTACTGTTGTTCAAACTGAAGATGAAATAGCAGCAGCAACGATGGCGA
TTGGAGCTAACTATGCTGGCGTACGAGGCTTTACAGCGAGTGCGGGTCCAG

40 GTCTTTCTTTAATGATGGAATCTATTGGATTGTCTGGTATGACTGAAACGC
CATTAGTCATTATTAATACTCAAAGAGGTGGTCCTTCTACTGGCTTACCAA
CAAAGCAAGAACAATCAGATTTAATGCAAATGATTTATGGTACCCATGGTG
ATATTCCGAAAATTGTCGTTGCTCCTACAGATGCTGAAGATGCGTTTTATC
TTACTATGGAAGCATTTAATTTAGCTGAAGAATACCAATGTCCAGTCATTC

45 TGTTAAGTGATTTACAATTATCATTAGGAAAACAAACTGTTAAAACACTCG
ATTATAATAAAATCGATATTCGTCGTGGAGAAATAATACAGTCAGATATCG
AGAGAGCTGAAGATGATAAAGCATACTTTAAAAGATATGCATTAACAGCTA
GTGGCGTATCACCACGACCAATACCAGGTGTTAAAGGTGGTATACATCATG
TAACA

50

Sequence 3473
step.1003f10.cons.ok
ATGGTACGTATGAAATGGCAAATACCCTAGCTGAAAACAATAATCCCTATG
55 CATATGTTGTCAGTGCATTAAGAAAAAACGTTAAAAGCGTGTATCGTGCCTT
TGAAAGAGGTGGAAAATCATGCTTAATCAATTTGTATGGGTTATTTATCCA
TATTTATGTTTAGCAATATTTGTTATTGGACATATCGCAAGATATAAATAT
GATCAATTTTCATGGACAGCAAAGTCGAGTGAAATGATTGAAAAGAAACGA
TTAAAATGGGGAAGTTTACTTTTTCATTTAGGTATTATTCCGGTATTCTTT

```
        GGTCATGTTGTGGGATTATTGATTCCCGCTAATTGGTTAGAGGCGATAGGA
        GTAAATAATCATATCTATCATATAGGTGCAGTTTATATAGGTAGTGTTTTT
        GGTATAATAACATTAATAGGAATGTTGTTGCTAACTTTAAGACGACTATCC
        ATCAAAAACGTTAGACGATTAAGTTCATTTTCAGATATATTTGTGAATATC
    5   GTTTTGTTGATTATTTTAATAATGGGTTGTTATTCTACGCTTGTAACCAAT
        GCGATTCAACCTGAATTTGATTATCGTCAAACCATTGCGATATGGTTTAGA
        CATTTATTCATGTTTTCTCCAAATGCTGACTTAATGTTAAACGTGCCTTGG
        TCGTTTAAACTGCACATATTATTAGGGTTTACAGTGTTTGCGTGTTGGCCA
        TTTACTCGTTTAGTACATGTTTGGAGTGTACCACTGTCTTATATGAACAGA
   10   AGATATATTGTTTATCGCAAAAACAAAATTTAATTTATTATGAGGTGAAAG
        TGTGGAACCAGAATCTTTAAAACATAATCAACATTTACAAATGAATTTAGA
        TAAGTTGAGAGCACAAGAGGGTTATGATTTTGGTGGTATCGCTTTATATGA
        TTATCATCACACTTCATCACCAATTAAATGGCAATATGTTTCAGGTAACAC
        AAATGATAGATATAAACTTATCATTTTGAGAAAGGGTAGAGGGCTTGCTGG
   15   AATGGTGATGAAAACCGGTAAGCGTATGGTTATTGCTGATGTAGATACAGC
        TTTATCTCCAGAAGAGAAAGTTAAATTTCCAATCATTCTTAGTGAGTCATT
        GACAGCTGTAGTTGCAGTCCCTTTATGGTTAGAAAATTCAATGTATGGCGT
        TTTATTATTAGGTCAAAGAAATCATCAGCCGTTACCTCAGTCATTGGACCA
        ACTTAATATTGAAAAACAAATCGGTATTTTTACAGAAATAAACTAGGTGGT
   20   AAATATGTTGGAGCAAACTGATTTAAGTTTAGAGCAATTACTTAAGAATTA
        TTATGAAACCACGAACGAGAAAATTGTATTTGTTAATAGACAAGGCAAAAT
        TATTGGTATGAATGACGCAGCAAAAGATATTTTAACTGAGGAAGATAATTA
        TAATGCTATGACAAATGCGATTTGTCATCGATGCGAAGGATACTCTAATGA
        ATATGATGTACAATCGTGTAAAGATTGTTTTTTTAGAGACAACGCAATTACA
   25   ACATTCCAATTTCCAAGTATTTATGAAGACAAAAGATAATGAAATTAAGCC
        TTTTACAGCTATGTATCAAAATATTGATGAACAAAGAGGTATTAGTGCATT
        TACCTTACAGAATGTGGCGCCTCAGATTGAAAGGCAAGAAAAAATGTATCA
        ACAAAAAATGTTACATCGTTCAATTCAAGCACAAGAAAATGAACGAAAGCG
        TATTTCTAGAGAATTACATGATAGTGTAATACAGGATATGCTCAATATAGA
   30   TGTTGAACTAAGGCTTTTGAAGTATAAGCACAGGGATAAGGTGTTAGCTGA
        AACATCTCAACGTATAGAAGGCTTATTATCACAGCTTATTGATGATATTAG
        AAATATGTCTGTTGAATTAAGACCTTCTTCTCTCGACGATTTAGGCATTGA
        AGCAGCTTTTAAATCATATTTTAAACAGTTTGAAGAAAATTATGGTATGCA
        TATTAAATATGATTCGAACATTAAAGGCATGCGTTTTGATAATGAAATTGA
   35   AACAGTTGTGTATCGTGTAGTTCAAGAGGGTGTATTTAATGCTCTAAAATA
        TGCTGAGGTTAATGAAATTGAGGTAAGTACGCATAGTGATGGCAAGCAGCT
        TGTAGCAGAGGTTGTGGATCGAGGTAAAGGGTTTAGTTTAGATCATCACCC
        TAAAGGCTCTGGACTTGGATTGTACGGAATGAGAGAACGTGCAGAATTAGT
        TAACGGTCATGTTAATATAGAGACACATATTAATAGAGGTACTATAATTAC
   40   ATTAGATATACCGATTTAAGTTGATGATGTAAGTTGGGGGAATTGAAGTGA
        AAATAGTTATAGCGGATGACCATGCAGTTGTTAGGACAGGATTTTCAATGA
        TATTAAATTATCAAGAAGATATGGAAGTTGTTGCAACTGCAGCTGACGGGG
        TTGAAGCTTATCAAAAAGTGTTAGAACATCGACCAGATGTTTTAATTTTAG
        ATTTGAGCATGCCGCCAGGAGAGTCAGGCTTAATCGCAACCAGTAAAATTT
   45   CTGAAAGTTTTCCTGATACTAAAATTTTAATACTTACGATGTTTGATGACG
        AAGAATATTTATTTCATGTGTTAAAAAGTGGTGCTAAAGGATACATTTTAA
        AAAATTCACCTGATGAGCAATTAATATTGGCCGTACGTACAGTATATCAAG
        GTGAAACTTATGTTGATATGAAATTGACGACGTCTTTAGTCAATGAGTTTG
        TCAATCAATCACAAACGGATGAAGTGTCATCATCTTCAGATCCATTTAAAA
   50   TTTTATCGAAACGAGAGTTAGAAATATTACCTCTTATAGCAAAAGGCTATG
        GCAATAAAGATATTGCAGAAAAGTTGTTTGTATCGGTGAAAACGGTAGAGG
        CACATAAAACGCATATTATGACGAAACTAAATTTAAAGAGTAAACCTGAAT
        TAGTTGAATATGCCTTAAAGAAAAAATTATTAGAATTTTAAATGTATATCA
        TGATTGATTATTAGGGCTTACTTAGTCTATGACTAAGTAGTCCTTTTTTTG
   55   TACCTTTATATAGGGAAAAATAAAAGTAAAATCAGGGTATGTCCTTATGTT
        CATATTTTCACAAACATATTAAAATGAATTTGTGATTGGTACTTTGTTTAT
        AATTTTGTTTTATATTATTAATTTTGACCTCTTAT
```

Sequence 3474
step.1003f11.cons.ok
TTCGTCATTGTTCGACTTTATTTGTTAATAAAATAGATAAGATAGAGGTGG
AAGAAACCGCCCGCTTGCTACGTCAACTCGAGCGTCTCAATAGCGATGCCA
ATATTCAAGTTGGTCAATTTGGAGAATTAAATTTAAAATCACTGCTAGAGC
CAACACATATAAATTCAAATGCATGTGGCACTTTGCATAGTAATATAAATC
ATCAATTCATCGAAAATCCTAGGCTACAAACAAAAGAAGAAATGATTAGTG
CGTTAGATAACTTGCCTCAAGATGTTTACCGTGTCAAAGGGTTTGTTCGTT
TTTCAGATCAGCAACACGTTTATTTAGTACAGTATGCACAAGGAAATATAG
AATTATCTCCCATTCAACTTAAAAACGATGTACCATTGTACCTCATTGTTA
TAGGAAAACATTTAAAACAAATÁCAATTTGATTTATAAAAACACATCTCTC
AATTATCATCAGTTAAACTTTATACATAAACAAACCTCTAAATGACTTTTG
TTCATTTAGAGGTTTTTGGTACTTATTTCTATTGTTCGTCTTTTATTAATG
TTTGCGCTTTTGAACTGATTCCATGCTTACTTCTACTCATTATTTACCTCA
CTTGGATTAAATTCCAAACTATCCGCAAAGTAGACGCTTTCTTCTTTATTG
TCGACAAAAGAAAACTCGATATCTTTATAACCTATAGATGAACCTTTAAGG
TCACCGTCCCCTTTTAATATTAGTTTAGGAGCTCTCTTAATCGGAATATCA
TACCTCTTTCTGAGTTGCTTCACATTGTCATCTTCATTACTTAATTGGTAC
TCTGCCGAATAGCTTGGTACGTTTGGGTTATAAGACACATTTCCATTCTTG
TAGTCTTTTAAATTTTTAAAATTCCCATATTGAGAGAAAAACTGAAAGTTT
TTAATTTCATTTTTTAATTTTTCATCGGCGATAGGTTTAGTCGGTTCAATT
TTATTGTTTTTAAGGCGAACCGGATATTCTTTATCTTTACTGTGCACTCTC
CCTTTTTTATCTTCTGTAATTATATTTGTATAAAAATGACCAGTCGTCTTT
CTAGTATTTCTATTCATATACAAAACCATACCTCTAGATTTCATCGCTTGA
TCTTTTTTCTGAATATTCATTTCTGAATTAATAATCCATGTCCCTTTGTCC
TCTTTTTCAAATTCTTCATCACGATAGCCTTCTTTATCGTATAAATCTTCT
AAATTTTTAATTGGATACATACTTAATGTTTTATTAAAATTTTCTTTAATT
TCCGCTTCTTTATTATTTTCTTTATTCATAAATCCACATCCGCCAATAAAA
AACGTTAATATTAAAATGCTTATGTAAGAAATGCACTTCCCATAATATTTC
ATCTTTTCACATCTCTCTATTATTGTATACATTTATTAAATATACATTTTA
TTGATATAAATATCAGTTTATAAAAAATTCAATTTTATATCATCAAATTATT
CATACACATTCAATTTAATTTATTTATATGTTTTTTTTCGGCGAAACAATCA
AAAAACCTCATTAAATGTACCGCAATCATACATCTAATGAGGCTATTATCG
TTCTTTTAATCGCTACTTACGTGTAATACCATCTAACTCTTCGACTTTCAA
CATATTATTTATCATACTCATATAGAATTGACCTAATTCTTCTCCAGTTCG
CTTATAATCAGTAGTATCTAAGCCGACTAAATCTAGATGATCCCAATCATA
TTGTACGGGTCTCACTTGCCAAATACCCTTGTCAGTTGCTAGGTTCATCAT
ACCTACCTTCTTAAATGCTTCATCACTTGGATGTAGTGAAGAAGACACAGG
TACTATGCCATCATTTACTCTGACATTTTTATTATCATCTCCACCTATCAC
ACGACTTGTTAAATCGAATAGTGGGAATTGTCTAATATTCGGCACTTCATT
GCCTAATGGTCCAGTATGTGTTGCAGCACCTGTATATGATGTATAGACGAT
ATTAGGATTCAATGTCGTCATTTGGTTTAACTTTTCTGCTCCAACAGTTGT
TAAATCATTTACAGCCTGATCTTCAGTCTCCCAAACTTTACTATTCGCTAT
ACGTTTTGCATATTCAGCATATGATTCATTAGGTTTCTGTTTGAAGCCCCA
TTGAGAAAAACCTAGTTCTAAATCGAGCGCTTTAGTTCCACCAATTTTTCC
AATTCTATTAATTGTATCTTTGATAAATTTAGTCGACCCTAGTTTATCTGC
AGCAGGTGTGCCATTATGAGGTGTTCCTAATGTAGTAATCGTAGACACCAT
GTTATCTTGTCCACCTTTAAACAAATCAGATACCGTACCACCATATTGACG
TTGATAATCTATTTCTTCTTGATTTCCATTTCTTAAAAAATGTTCCATCAA
GCGTATCGTTTGGCCACCCATACTATGTCCAACAAGATGTATCTTTTTACC
TGGTTCCCAATCAGGCATGATGCCTTCATATGTTCTGCCATAACGCTTGTG
ACCATATTTTGCAGCATGTGCTGCACCATAATCTACTCTTCCACCTTTAAT
ATAATAATACAGTTCAACAGCACGGTCATAATTGCTGCTAAATGCTCCTAC
ATTGGCTTCGTGAACTCGGTAACCTAATTTTGTAAGTTCTTGTTTCACGTT
ATATTTAGTACCACCCCAATAATTTGGGTACATGCTGAATGAATCTTCACC
GACTAAACCTACAAATCCATGTACAAACACGACTGGATATTGATTTTTATA
TTGAGCTTGCGCAGTAAGTTGATTGATATGTTTAGTTTGTTTTGATTTCAC
AACGCGTGGTGAAGTTAATGATGTTTGCTTTTTTTTGATTTGTTTTAACTGT
TTTATCAGTTTTTGTATTACCTATTTTGTGTACTTTAGAAGATGGTTGTTT

TACTGTTTTATCTCTATCTCCTATAGACTGTGATTGTTTAGGTTGTGTTCC
TTTTGAAGATGGGGCTTGAGTTTGATCAGATGGGTGTTGCTTTTTATCTGA
TGAATCTTGATCATTATTTTGGTTTTGAGTATTAGGTTGATGTTCTGTTTG
TGAGTGTTGATTATCTTTAAACTGTTGTTTGCTTTATCATCTTGATTTTG
GCCTTTTGGATATGTATCCTGACGTGAATTTGGGGTAACATGATTGTCTTT
TGATTGTTGCCCCTTTAAAGTTTGTTCTTGGTCATTTAATGGTGACGTTTG
TGTTGAATCTTGTTGAGTTTGATTATTCTTTTTTTGTACATTTTCATCAGA
TTGGTTCTCTTGATTAGTCTTTGATTGGCTAGCTGTGTCACCCTCTTGATT
CAATTCGTATGATAAATGAGTTGTATCTTCTTTGTGTTCAACTTGTTCTGT
TTGTTGTGACTGTTCGTTAAAATCTGATGATGATTGTGTCATTTCAGCAGC
TTGAGC

Sequence 3475
step.1003f12.cons.ok
AACGAAATACAAGAAAATGTGACTGAAAAGACGAAAGGTCATAAATTAAAA
AAAAGTGCAGCTAAAACAACAGCCCTAGTTGGTGGAGCATTTACATTTAAT
ATGTTGAATAATCATCAAGCATTAGCTGCCTCAGAAACACCAATCACCTCT
GAAATTTCATCCAATAGCGAGACAGTAGCCAATCAAAATTCAACTACGATT
AAAAACTCACAAAAAGAAACAGTCAATTCTACAAGTTTGGAATCTAACCAT
TCTAATAGTACAAATAAGCAAATGTCTTCAGAAGTTACAAATACAGCTCAA
TCAAGTGAAAAAGCTGGAATTAGTCAACAAAGTAGTGAAACATCAAATCAA
TCATCTAAGTTAAATACATATGCCTCCACAGACCATGTAGAGAGTACAACT
ACTAACAATGATAATACTGCACAACAAGATCAAAATAAGTCTTCGAATGTA
ACCTCTAAGTCAACACAATCAAACACGTCATCCTCAGAAAAGAGCATTAGC
TCCAATTTAACCCAGTCAATCGAAACAAAAGCAACCGATTCATTAGCGACC
AGTGAAGTACGTACTAGTACAAATCAAATATCTAATCTGACATCAACATCT
ACTTCAAATCAATCGAGTCCTACTTCTTTTGCAAATTTAAGAACATTTAGT
AGATTTACCGTTTTGAATACGATGGCAGCACCGACAACAACGTCCACGACA
ACAACTTCAAGTCTGACATCTAATTCTGTTGTGGTGAACAAAGATAACTTT
AATGAACATATGAATCTATCTGGATCTGCGACGTATGATCCAAAAACAGGT
ATTGCTACCTTAACGCCAGATGCATATAGTCAAAAGGGTGCCATATCTTTA
AACACTCGATTAGATTCAAACCGTAGCTTCCGTTTTACAGGTAAAGTTAAC
CTTGGTAATAGATATGAAGGTTATTCTCCTGATGGTGTAGTAGGTGGAGAT
GGCATTGGCTTTGCATTTTCACCAGGTCCTTTAGGACAGATAGGTAAAGAG
GGGGCTGCCGTTGGAATAGGTGGTTTGAATAATGCCTTTGGATTTAAATTG
GATACGTATCATAACACATCACCTCCTAAATCTGATGCTAAAGCAAAAGCA
GACCCACGTAATGTTGGTGGTGGTGGCGCTTTTGGTGCCTTCGTAAGTACA
GATAGAAATGGTATGGCTACCACTGAGGCATCATCTGCGGCTAAATTAAAT
GTACAACCTACTGACAATTCATTCCAAGATTTTGTCATTGACTATAATGGT
GATACAAAAGTGATGACAGTGACGTACGCTGGACAAACTTTTTACGAGAAAT
CTTACAGATTGGATTAAAAACAGTGGTGGTACGACGTTTTCTCTATCTATG
ACTGCCTCAACTGGTGGCGCAAAAAAACTTACAACAAGTTCAATTTGGAACA
TTCGAGTATACAGAATCAGCTGTTGCTAAAGTACGCTATGTAGATGCAAAT
ACTGGTAAGGATATTATTCCACCTAAAACCATTGCAGGTGAAGTTGATGCG
ACTGTGAATATAGATAAACAATTAAACAACTTGAAAAATTCAGGTTACAGT
TATGTTAGTACAGACGCTTTACAAAACTCCAATTATTCAGAAACATCAGGT
ACACCTACACTTAAATTAACTAACTCAAGCCAAACGGTGATTTATAAATTC
AAAGATGTTCAAGGTCCTCAAATTAGTGTTGATAGTCAAACTAGAGAAGTT
GGAAAGACCATTAATCCAATTACAATTACTACAACTGACAATAGTAAAGAC
GTATTAACTACAACTGTGACAGGTCTACCTTCAGGGTTATCTTTTGATCAA
ACGACAAATACAATTACTGGCACGCCAAGTGAAGTAGGAACTACAACTGTG
ACAGTTAATACTACTGATGCTACTGGGAACGTAACATCTAAGCAATTTACA
ATAACGATTCAAGATACAATCAGCCCTGTTGTAAATGTGACGCCAAGTCAA
GCATCAGAGGTCAACGACAACAGCTATAAAGAACCATAAAATATTACTTAT
TCCACCTAGTACAGCGACAATAGGACCTCCATGAGCAACTTTATCACCGAC
ACCTCCAACCGCCGCTATTACTGAAGCAATCATCGCGCCAATCATATTGGC
TGGAATGATTCTTAGAGGATCTTGTGCAGCAAAAGGGATTGCCCCTTCAGT
AATACCAAATAATCCCATCGTAAATGATGCCTTACCCATTTCTTGTTCTGC

```
       TTTATTGAATTGGTGTTTACGAACAAATGTAGCTAAACCTAAACCAATCGG
       TGGTGTACATACTGCTACAGCAACCATTCCCATCACAGTGTAGTTGCCTTC
       AGCAATTAATGCAGAACCGAATAAGAATGCTACTTTGTTTACTGGACCTCC
       CATATCAAAAGCAATCATCGCGCCAATAATAAGAGCAAGAATAATGATGTT
  5    AGCACCTTGCATTCCTTTTAACCATGATGTTAATGCACCAAATATATTTGA
       AATTGGTGCGCCTATTACAAATATAAAAATGAGACCAACTATTAAAGAAGA
       TAGAATAGGTATAATAATAATAGGCATAATAGGAGCCATAGCTTTAGGAAC
       TTTAATCTGTTTAATCCATTTTGCAATATAGCCCGCTAAGAAACCTGCGAC
       GATACCACCAAGGAAACCGGCTCCTGCTTCACTTCCATATAAACTACCATC
  10   AGCAGCAATGGCACCACCAATCATACCTGGAACAAGACCAGGCTTATCAGC
       AATACTCACCGCGATATAACCAGCAAGGATGGGAACCATGAATTTAAACGA
       TAAACTACCAATATTTTCAATAGATTTCCAAAATGAATCTTCGGGGATAAC
       TAATCCTTTTGGAGTGGTGTGTCCTCCAAGAGTTAAGGCAATAGCTATGAG
       TAAACCGCCAACGACGATAAATGGAACCATGAAGGAAACACCATTCATTAA
  15   ATGCTGATATACCATTTGAATATTACTTTTATGAGATT


       Sequence 3476
       step.1003g01.cons.ok
  20   CGCCAAAACTGATATTGAATGGACATGTTGATGTAGCTTCTGTAGATGACG
       ATCAATATTGGCAGTATCCACCTTTTAAACTTACCAACAAAGAAGAATGGT
       TATACGGTCGTGGCGTTAGCGATATGAAAGGTGGTATGTCTTCATTATTCT
       ACGTCTTGGAGCAATTACATCAAGAGGGGCAACGTCCAGAAGGTGATATTA
       TTGTTCAATCAGTAGTCGGTGAAGAAGTAGGTGAAGCAGGAACTAAACGTG
  25   CATGTGAAATAGGACCTAAAGGTGACTTAGCCCTTGTCTTAGATACGAGTG
       AGAATCAAGCACTTGGGCAAGGTGGCGTGATTACCGGATGGATTACAGTTA
       AAAGTAAAAATACAATACATGATGGTGCGCGTAGTCAAACGATACATGCTG
       GTGGGGGCTTGTTTGGTGCAAGTGCCATTGAAAAAATGACAAAGGTGATTC
       AATCGCTTAATGAACTTGAAAGGCATTGGGCTGTCATGAAGAAGAGCCCTG
  30   GAATGCCTCCAGGTGCGAATACAATTAACCCAGCTGTCATAGAAGGTGGAC
       GTCACCCTGCATTTATTGCAGATGAATGTCGATTATGGATTACTGTTCATT
       ACTTACCGAACGAAAGTTATGAATCTGTAGTTAATGAAATAGAGCAATATT
       TAAATAAGGTTGCAGAAGCAGATGTATGGCTCAGAGAGAATCCACTTGAAT
       TTGAATGGGGTGGTACATCCATGATTGAGGATAAAGGAGAAATCTTCCCAA
  35   GTTTCACTGTTCCGACACATCATCCAGGTTTTAAGCAATTAGAAGAAGCAC
       ATGAACATATTCATAATAAAAAGCTTGAACATGGTATGAGTACAACTGTAA
       CTGATGGAGGTTGGACAGCACATTTTGGCATTCCCACGATATTATATGGCC
       CAGGTAGTTTAGAAGAGGCACATAGTGTAGATGAGAAAATAAAAGCAAAGG
       AATTAGCTCAATATAGTGATGTTTTATATACATTTTTAAAAGAGTGGTATG
  40   CACACCCACAATCCTATAAATCATCATAGATAAAAAAGAGGTACAAGCACG
       ACCTTTTATACTCACAAAATGTTGAGTTAAAGTGATGTGTTCTGTACCTCT
       TAATTATTATAAAGAATAACAAGTTGTATATTTATAAAAAGATATTTATTA
       ATTTCGTGTTCATTTCATATAAAATACAATTGTCAAGGTGAGTTTATCTTT
       TAAATGTACTTTTTAATTTTCGTATCCAACTACCACCACGTCTAGAATGAA
  45   CGGGTTGTCCAGAAGGTTGTACCGATTGTTTTTTTAGTTCTTCTCCATACT
       TAATTGCTTCTTGCATCAAGTAAGATTGAGAGTTTAAAGGAGATGAATTGT
       TTTCGACATAATGATAAACGCCTTGTGCATCTTGGTAACGACCTTTTTGGT
       TGTCAGATAATTGTAAATTCATATAGTTAACTAATCGTTGTCCTATTGATT
       TATCTTCAACAGGGAATAATATTTCGACACGTTTAATCATATTACGTGTCA
  50   TATCGTCAGCTGAAGATAAATAAATATGCGCCTCACCATTATTATGGAAGT
       AGTAAACGTGAATGTTCAAGCAAACGACCCACAATACTAACAACCTCTA
       TATTTTCGCTAATACCTGGAATGCCTGGTTTAAGACAACATATACCACGAA
       TGATGAGTTGTATTTTAACGCCTGCTTGGGATGCTTCGAAGAGCTTTTCGA
       TAATCGTTTTATCGGTTAAAGAGTTCATTTTCATCATAATTTTACCGTTAC
  55   CATGTTGTAAATGACTACGTATTTCTTTATCGATACGATCAATGAAGACGT
       CTCGAATATCGTATGGTGCTACAATCAATTTATTGTATTCTGGTTTTGTTG
       AGTAACCACTCAAGTAATTAAAGAAGTTAATTGCATCCTCAGCGATATCTT
       TATTTGTCGTGATGATACCCATATCTGTGTATAATTTAGCAGTTTTATCGT
       TATAGTTACCTGTGCCTAAATGAACAAATGACGTAAGTTTATTGTTGATGC
```